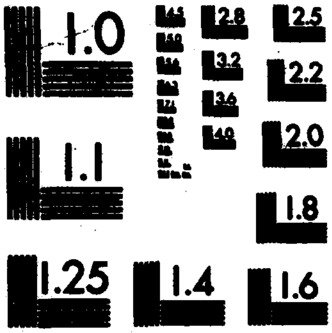
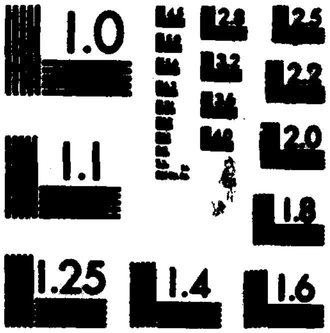
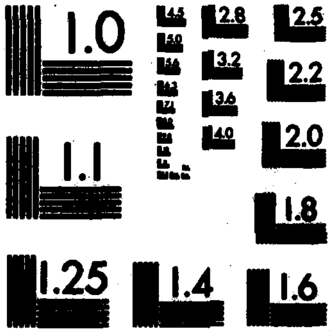
AD-A120 811 PROCEEDINGS OF THE 1982 ARMY SCIENCE CONFERENCE HELD AT 1/6
THE UNITED STATES..(U) DEPUTY CHIEF OF STAFF FOR
RESEARCH DEVELOPMENT AND ACQUISITIO.. 18 JUN 82
UNCLASSIFIED F/G 5/2 NL



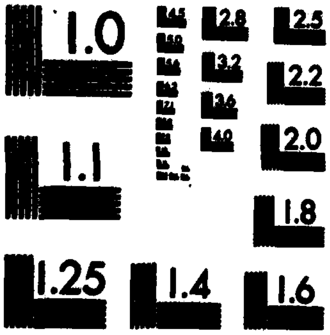
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



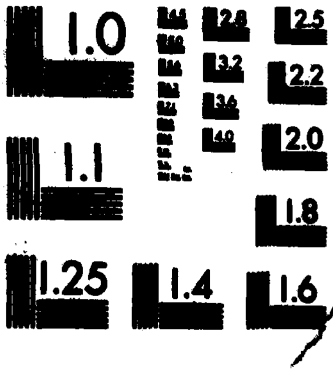
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

DEPARTMENT OF THE ARMY
OFFICE OF THE DEPUTY CHIEF OF STAFF
FOR RESEARCH, DEVELOPMENT, AND ACQUISITION
WASHINGTON, D.C. 20310

REPLY TO
ATTENTION OF

DAMA-ARZ-D

18 June 1982

SUBJECT: Proceedings of the 1982 Army Science Conference

SEE DISTRIBUTION

1. The thirteenth in a series of Army Science Conferences was held at the United States Military Academy, 15-18 June 1982. The conference presented a cross section of the many significant scientific and engineering programs carried out by the Department of the Army and provided an opportunity for Department of the Army civilian and military scientists and engineers to present the results of their research and development efforts before a distinguished and critical audience.

2. These Proceedings of the 1982 Army Science Conference are a compilation of all papers presented at the conference and the supplemental papers that were submitted. The Proceedings consist of four volumes, with Volumes I through III unclassified, and Volume IV classified.

3. Our purpose for soliciting these papers was:

a. To stimulate the involvement of scientific and engineering talent within the Department of the Army.

b. To demonstrate Army competence in research and development.

c. To provide a forum wherein Army personnel can demonstrate the full scope and depth of their current projects.

d. To promote the interchange of ideas among members of the Army scientific and engineering community.

4. The information contained in these volumes will be of benefit to those who attended the conference and to others interested in Army research and development. It is requested that these Proceedings be placed in technical libraries where they will be available for reference.

James H Merryman

JAMES H. MERRYMAN
Lieutenant General, GS
Deputy Chief of Staff for Research,
Development, and Acquisition

DISTRIBUTION:
Office, Under Secretary of Defense for Rsch & Engineering, Wash, DC 20310
HQDA, ODCSRDA, ATTN: DAMA-ARZ-A, Washington, DC 20310
HQDA, ODCSRDA, ATTN: DAMA-ARZ-D, Washington, DC 20310
HQDA, Deputy Chief of Staff for Logistics, Washington, DC 20310
HQDA, Deputy Chief of Staff for Operations, Washington, DC 20310
HQDA, Deputy Chief of Staff for Personnel, Washington, DC 20310
HQDA, Asst Chief of Staff for Automation and Communication, Wash, DC 20310
HQDA, Assistant Chief of Staff for Intelligence, Washington, DC 20310
HQDA, Office of the Chief of Public Affairs, Washington, DC 20310
Office of the Surgeon General, ATTN: DASG-RDZ, Washington, DC 20310
Office, Chief of Engineers, ATTN: DAEN-RDZ-A, Washington, DC 20314
Office, Chief of Engineers, ATTN: DAEN-RDZ-B, Washington, DC 20314
Office, Chief of Engineers, ATTN: DAEN-RDM, Washington, DC 20314
COMMANDERS/DIRECTORS
US Army Ballistic Missile Defense Systems Command, Huntsville, AL 35807
US Army Computer Systems Command, Ft. Belvoir, VA 22060
US Army Concepts Analysis Agency, Bethesda, MD
US Army Operational Test and Evaluation Agency, Falls Church, VA 22041
US Army Mat Dev & Readiness Cmd, ATTN: DRCDRA-ST, Alexandria, VA 22333
US Army Mat Dev & Readiness Cmd, ATTN: DRCLD, Alexandria, VA 22333
US Army Armament R&D Cmd, Dover, NJ 07801
Ballistic Research Lab, Aberdeen Proving Ground, MD 21005
Chemical Systems Lab, Aberdeen Proving Ground, MD 21010
Large Caliber Weapons Systems Lab, Dover, NJ 07801
Fire Control & Small Caliber Weapons Systems Lab, Dover, NJ 07801
US Army Aviation R&D Cmd, St. Louis, MO 63166
USARTL, Ames Research Center, Moffett Field, CA 94035
USARTL, Aeromechanics Lab, Moffett Field, CA 94035
USARTL, Applied Technology Lab, Ft. Eustis, VA 23604
USARTL, Propulsion Lab, Cleveland, OH 44135
USARTL, Structures Lab, Hampton, VA 22665
US Army Avionics R&D Activity, Ft. Monmouth, NJ 077
US Army Aviation Engineering Flight Activity, Edwar [illegible] CA 93523
US Army Communications & Electronics Cmd, Ft. Monmouth [illegible] 7703
Center for Communications Systems, Ft. Monmouth, NJ 07703
Center for Tactical Computer Systems, Ft. Monmouth, NJ 07703
Center for Systems Engineering & Integration, Ft. Monmouth, NJ 07703
US Army Electronics R&D Command, Adelphi, MD 20783
Atmospheric Sciences Lab, White Sands Missile Range, NM 88002
Combat Surveillance & Target Acquisition Lab, Ft. Monmouth, NJ 07703
Electronics Technology & Devices Lab, Ft. Monmouth, NJ 07703
Electronics Warfare Lab, Ft. Monmouth, NJ 07703
Ofc of Missile Electronic Warfare, White Sands Missile Range, NM 88002
Harry Diamond Laboratories, Adelphi, MD 20783
Night Vision & Electro-Optics Lab, Ft. Belvoir, VA 22060
Signals Warfare Lab, Vint Hill Farms Station, VA 22186
US Army Missile Command, Redstone Arsenal, AL 35898
US Army Missile Lab, Redstone Arsenal, AL 35898
US Army Mobility Equipment R&D Command, Ft. Belvoir, VA 22060

US Army Natick Laboratories, Natick, MA 01760
US Army Tank-Automotive Command, Warren, MI 48090
 Tank-Automotive Systems Lab, Warren, MI 48090
 Tank-Automotive Concepts Lab, Warren, MI 48090
US Army Test & Evaluation Command, Aberdeen Proving Ground, MD 21005
 US Army Aberdeen Proving Ground, Aberdeen Proving Gnd, MD 21005
 US Army Dugway Proving Ground, Dugway, UT 84022
 US Army Electronic Proving Ground, Ft. Huachuca, AZ 85613
 US Army Tropic Test Center, APO Miami 34004
 US Army White Sands Missile Range, White Sands Missile Range, NM 88002
 US Army Yuma Proving Ground, Yuma, AZ 85364
US Army Research Office, Research Triangle Park, NC 27709
US Army Materials & Mechanics Research Center, Watertown, MA 02172
US Army Human Engineering Lab, Aberdeen Proving Ground, MD 21005
US Army Mat Systems Analysis Activity, Aberdeen Proving Ground, MD 21005
US Army Foreign Science & Tech Center, Charlottesville, VA 22901
US Army Research, Development & Standardization Group (Europe)
US Army Satellite Communications Agency, ATTN: DRCPM-SC-11, Fort Monmouth, NJ 07703
US Army Training & Doctrine Cmd, ATTN: ATCD, Ft. Monroe, VA 23651
US Army Health Services Command, Ft. Sam Houston, TX 78234
Institute of Surgical Research, Ft. Sam Houston, TX 78234
US Army Medical R&D Command, Fort Detrick, Frederick, MD 21701
 Aeromedical Research Lab, Ft. Rucker, AL 36362
 Institute of Dental Research, WRAMC, Washington, DC 20012
 Letterman Army Inst of Research, Presidio of San Francisco, CA 94129
 Medical Bioengineering R&D Lab, Frederick, MD 21701
 Medical Rsch Inst of Chemical Defense, Aberdeen Proving Gnd, MD 21010
 Medical Rsch Inst of Environmental Medicine, Natick, MA 01760
 Medical Rsch Inst of Infectious Diseases, Frederick, MD 21701
 Walter Reed Army Inst of Research, Washington, DC 20012
  Walter Reed Army Medical Center, Washington, DC
Armed Forces Radiobiology Rsch Inst, Bethesda, MD 20814
US Army Cold Regions Rsch & Eng Lab, Hanover, NH
US Army Construction Eng Rsch Lab, Champaign, IL
US Army Engineer Topographic Labs, Ft. Belvoir, VA 22060
US Army Engineer Waterways Experiment Station, Vicksburg, MS 39180
US Army Rsch Inst for the Behavioral & Social Sciences, Alex, VA 22333
 ARI Field Unit, Ft. Benjamin Harrison, IN 46216
 ARI Field Unit, Ft. Sill, OK 73503
 ARI Field Unit, Presidio of Monterey, CA 93940

COMMANDANTS:
Academy of Health Sciences, ATTN: HSHA-CDM, Ft. Sam Houston, TX 78234
US Army Air Defense School, ATTN: ATSA-CDM-A, Ft. Bliss, TX 79916
US Army Armor Center, ATTN: ATZK-CD-SD, Ft. Knox, KY 40121
US Army Chemical School, ATTN: ATZN-CM-CS, Ft. McClellan, AL 36205
US Army Engineer School, ATTN: ATZA-CD, Ft. Belvoir, VA 22060
US Army Field Artillery School, Ft. Sill, OK 73503
US Army Infantry School, Ft. Benning, GA 31905
US Army Intelligence Center and School, Ft. Huachuca, AZ 85613

US Army Logistics Center, Ft. Lee, VA 23801
US Army Ordnance Center and School, Aberdeen Proving Gnd, MD 21005
US Army Signal Center, Ft. Gordon, GA 30905
US Army Transportation School, ATTN: ATSP-CD-CS, Ft. Eustis, VA 23604
SUPERINTENDENT:
US Army Military Academy, ATTN: Technical Library, West Point, NY 10996
COPIES FURNISHED:
Defense Advanced Research Projects Agency, Arlington, VA 22209
Defense Technical Information Center, Alexandria, VA 22209
The Army Library, ATTN: ANRAL, Washington, DC
Chief, US Army Field Office, HQ AFSC/SDOA, Andrews AFB, MD 20331
Cdr, HQ Fort Huachuca, ATTN: Tech Reference Div, Ft. Huachuca, AZ 85613
Cdr, 1st COSCOM, HHC/SOTI, Ft. Bragg, NC 28303
Naval Air Systems Command (Code 310-A), Washington, DC 20361
Naval Research Library (Code 2627), Washington, DC 20361
Office of Naval Research (Code 102), Arlington, VA 22217
HQ US Marine Corps (Code RD-1), Washington, DC 20380
Air Force Systems Command, Andrews AFB, Washington, DC 20331
Lawrence Livermore Lab, Univ of California, Livermore, CA 94550
Los Alamos Scientific Lab, Los Alamos, NM 87544
Southwest Research Institute, San Antonio, TX 78228
United Nations Library, New York, NY 10017

PROCEEDINGS
OF THE
1982 ARMY SCIENCE CONFERENCE

UNITED STATES MILITARY ACADEMY
WEST POINT, NEW YORK
15-18 JUNE 1982

VOLUME I

Principal Authors A through G

Accession For
NTIS GRA&I ☒
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A

DTIC
COPY
INSPECTED
2

TABLE OF CONTENTS

PROCEEDINGS OF THE 1982 ARMY SCIENCE CONFERENCE

Author	Title	Vol	Page
Adams, Nelson L.	See Rickett, Daniel L.	3	117
Albright, Ronald	See Patton, John F.	3	1
Albritton, Gayle E.	See Cost, Van T.	1	267
Alexander, Donald	See Patton, John F.	3	1
Allan, Barry D.	✓A Gelled Oxidizer for Tactical Missiles	4	1
Allan, Barry D.	See Ayers, O. E.	1	17
Alster, Jack	See Gilbert, Everett E.	4	87
Althouse, Mark L.	See Cox, Christopher S.	4	61
Antle, Lloyd G.	See Goicoechea, Ambrose	1	475
Ashley, Paul R.	See Duthie, Joseph G.	1	341
Ashman, William P. Lewis, James H. Poziomek, Edward J.	✓A Decision Tree for Chemical Detection Application	1	1
Atchison, Valerie L.	See Voorhees, James W.	3	425
AuCoin, Thomas R.	See Malik, Roger J.	2	309
Ayers, O. E. Allan, B. D. Melvin, W. S. Murfree, J. A. Wharton, W. W. Marcucci, P. J.	✓Laser Induced Polymerization Reactions in Solid Propellant Binders	1	17
Baba, Anthony J. Share, Stewart	✓Thermal Radiation Effects on Fiber Optics and Conventional Metallic Communications Cable	4	15
Babbitt, Richard W.	See Borowick, John	1	113

Author	Title	Vol	Page
Bagby, D. Gordon	✓ The Engineer Modeling Study	1	31
Barditch, Irving F.	See Cox, Christopher S.	4	61
Barr, Thomas A., Jr. McKnight, W. B.	✓ A New Midinfrared Laser in Hydrogen	1	47
Barsam, Helena F.	See Simutis, Zita M.	3	273
Basso, Michael J.	See McCreery, M. J.	2	357
Bates, Calvin	See Fischer, Paul	1	399
Baussus von Luetzow, H.	✓ Gravity Vector Determination from Inertial and Auxiliary Data and Potential Utilization of Generated Vector Component Information	1	61
Bayha, William	See Borowick, John	1	113
Beatrice, Edwin S.	See Lund, David J.	2	279
Beatrice, Edwin S.	See Randolph, David I.	3	95
Beatrice, Edwin S.	See Schuschereba, Stephen	3	173
Beatrice, Edwin S.	See Stuck, Bruce E.	3	371
Beatrice, Edwin S.	See Zwick, H.	3	449
Bedynek, Julius	See Patton, John F.	3	1
Berg, Richard	See Lunardini, Virgil J.	2	263
Berkhimer, Karl	See Sturdivan, Larry M.	4	209
Bertin, John J. Goodyear, Richard L.	✓ Multiple Launch Rocket System (MLRS) Fuze	1	77
Bexon, Roy	See Sturdivan, Larry M.	4	209
Bhatt, Ramakrishna T.	✓ Thermal Fatigue Behavior of FP Alumina/Magnesium Composites	1	89
Bieberich, M. J.	See Singler, R. E.	3	297
Binder, Michael	See Gilman, Sol	1	447

Author	Title	Vol	Page
Bingham, Gene J.	✓The Aerodynamic Influences of Rotor Blade Taper, Twist, Airfoils and Solidity on Hover and Forward Flight Performance	1	99
Binn, Leonard N.	See Lemon, Stanley M.	2	249
Black, Edward D.	See Porter, William L.	3	31
Bloom, Kenneth R.	See Zwick, H.	3	449
Borowick, John Bayha, William Stern, Richard A. Babbitt, Richard W.	✓Dielectric Waveguide Antenna	1	113
Bowden, Charles M.	✓ MICOM Program in Light Control by Light	1	125
Brandt, W. E.	See Henchal, Erik A.	2	61
Brodman, Bruce W. Devine, Michael P.	✓Microbial Attack of Nitrocellulose	1	141
Brown, Douglas R.	✓An Empirical Model for Near Millimeter Wave Snow Extinction and Backscatter	1	147
Buser, R. G.	See Rohde, R. S.	4	177
Bushell, M. Manriquez, R. Merkel, G. Scharf, W. D.	✓Aurora Flash X-Ray Facility as a Source-Region EMP Simulator	1	159
Campana, Joseph E.	See Friedman, Joseph E.	1	415
Campbell, Donn V.	Control of Parasitic Currents on Radiating Systems	1	175
Campi, Morris	Design of Microstrip Linear Array Antennas by Computer	1	187
Carchedi, Steven	See Groff, John N.	4	97
Caslavsky, Jaroslav L. Viechnicki, Dennis J.	Resolution of Factors Responsible for Difficulty in Growing Single Crystals of YAG	1	201

Author	Title	Vol	Page
Chandra, S.	See Rohde, R. S.	4	177
Chapin, Charles C. Conley, James H. Jamison, Robert G.	Generation of a System Which Exhibits an Isopycnic in the Conversion of Army Vehicles to Silicone Brake Fluid	1	209
Chu, Shih C. Steiner, James	A More Rational Approach to the Stress Analysis of Projectiles	1	225
Chubb, Douglas W.J.	Natural Language Translation in a Noisy Domain Using Occam's Razor as a Control Paradigm	4	31
Cohn, Stephen L. Pena, Ricardo	Munition Expenditure Model Verification: KWIK Phase I	1	239
Collins, John G. Link, Lewis E.	A Quantitative Evaluation of Groundwater Resources in Southwest Asia	4	47
Comminge, Hubert	Hardening of Armored Vehicle Suspension Components	1	255
Conklin, James J.	See McCreery, M. J.	2	357
Conley, James H.	See Chapin, Charles C.	1	209
Cost, Van T. Albritton, Gayle E.	Response of MX Horizontal Shelter Models to Static and Dynamic Loading	1	267
Cox, Christopher S. Althouse, Mark L. Barditch, Irving F.	Point and Remote Detection of Biological Aerosols Using Infra-Red	4	61
Crumley, Lloyd M. Schwalm, Robert C.	Modeling Weapons Crew Performance	1	283
Davenport, Otis A.	CCM Alternatives to Soviet Lasers Targeted Against the Helicopter Force	4	77
DeBellis, William B.	See Poston, Alan M.	3	47
DeLuca, Eugenio	See Prifti, Joseph J.	4	139
Deome, A. J.	See Singler, R. E.	3	297
DeVenuto, Frank Zegna, Angelo I.	The Development and Evaluation of Hemoglobin Solution as a Blood Substitute	1	299

Author	Title	Vol	Page
DeVenuto, Frank Zegna, Angelo I.	Field Production of Purified Sterile Water from Available Water Sources by Using a Portable Apparatus	1	315
Devine, Michael P.	See Brodman, Bruce W.	1	141
Dixon, R. E.	See Rohde, R. S.	4	177
Drolet, Anne M.	See Porter, William L.	3	31
DuBois, John	See Powanda, Michael C.	3	63
Dunn, D. A.	See Singler, R. E.	3	297
Dunn, William P.	A New Analysis of the Unwinding Ribbon as a Delayed Arming Device	1	325
Duthie, Joseph G. Ashley, Paul R. Upatnieks, Juris Liu, H. K.	Photonic Seeker Development	1	341
Elder, Alexander S. Walbert, James N. Zimmerman, Kathleen L.	Stationary and Traveling Loads in a Hollow Cylinder	1	353
Emery, Clarence E.	See Wolfe, Alan D.	3	435
Espeland, Richard H.	See Schwering, Felix K.	3	201
Essenwanger, Oskar M.	Turbulence Analysis by Use of the Fast Fourier Transform	1	367
Figucia, F. Williams, C. Kirkwood, B. Koza, W.	Mechanisms of Improved Ballistic Fabric Performance	1	383
Fischer, Paul Bates, Calvin Hartley, Joseph	A Novel Beam Bunching Concept for Millimeter Wave Tubes	1	399
Friedman, Melvin H. Campana, Joseph E. Yergey, Alfred L.	New Viewpoints in Mass Filter Design	1	415
Gall, Kenneth J.	See Rickett, Daniel L.	3	117

Author	Title	Vol	Page
Garinther, John M.	See Poston, Alan M.	3	47
Garn, Lynn E.	See Sharp, Edward J.	3	227
Garnett, Lamont W.	See Squire, Walter H.	3	341
Garrett, Paul F., Jr.	See Poston, Alan M.	3	47
Gauss, Arthur, Jr.	A New Type of EM Wave Absorbing Coating	1	431
Gentry, M. K.	See Henchal, Erik A.	2	61
Gilbert, Everett E. Sollott, Gilbert P. Alster, Jack Sandus, Oscar Slagg, Norman	Toward More Powerful Explosives--Polynitro Polyhedranes	4	87
Gilman, Sol Wade, William, Jr. Binder, Michael	High Energy Sulfuryl Chloride Batteries	1	447
Gladen, Curtis L. Link, Lewis E.	Thermal Camouflage of Fixed Installations: Project TIREX	1	459
Goicoechea, Ambrose Krouse, Michael R. Antle, Lloyd G.	An Approach to Risk and Uncertainty in Benefit-Cost Analysis of Water Resources Projects	1	475
Goicoechea, Ambrose Krouse, Michael R.	A Computer Based Interactive Model for Industrial Land Use Forecasting	1	489
Goldman, Ralph F.	Microclimate Cooling for Combat Vehicle Crewmen	1	503
Goodrick, Thomas F.	Development of Methods for Assessment of Gliding Parachute Applications	1	517
Goodyear, Richard L.	See Bertin, John J.	1	77
Gregory, Frederick H. Pearson, Richard J.	Analytical and Experimental Studies of the Response of a Cylinder to Nuclear Thermal/Blast Loads	1	531
Griffith, James R.	See Klinger, Liliana	2	223

Author	Title	Vol	Page
Groff, John N. Mirabelle, Rosemary M. Carchedi, Steven	Armored Combat Vehicle Technology Using the HITPRO/DELACC Methodology	4	97
Groves, Michael G.	See Twartz, John C.	3	411
Hagman, Joseph D.	Maintaining Motor Skill Performance	2	1
Hardy, G. David, Jr.	See Hiller, Jack H.	2	75
Harley, Samuel F.	Data Compression for Transient Measurements	2	17
Harris, Paul Presles, Henri-Noel	The Shock Front Rise Time in Water	2	33
Hartley, Joseph	See Fischer, Paul	1	399
Heavey, Karen R.	See Nietubicz, Charles J.	2	425
Heberlein, David C.	Detonation of Rapidly Dispersed Powders in Air	2	47
Heise, Carl J.	See Mando, Michael A.	4	111
Henchal, Erik A. McCown, J. M. Gentry, M. K. Brandt, W. E.	Rapid Identification of Dengue Virus Serotypes Using Monoclonal Antibodies in an Indirect Immunofluorescence Test	2	61
Herren, Kenneth A.	See Johnson, John L.	2	199
Hiller, Jack H. Hardy, G. David Meliza, Larry L.	Design of a Small Unit Drill Training System	2	75
Hoidale, Glenn B.	See Walters, Donald L.	4	239
Hollenbaugh, D. D.	Quantification of Helicopter Vibration Ride Quality Using Absorbed Power Measurements	2	91
Howe, Philip M. Kiwan, Abdul R.	A Theoretical Study of the Propagation of a Mass Detonation	2	109
Hsiesh, Jen-Shu	See McCreery, M. J.	2	357
Hubbard, Roger W. Mager, Milton Kerstein, Morris	Water as a Tactical Weapon: A Doctrine for Preventing Heat Casualties	2	125

Author	Title	Vol	Page
Hutchings, Thomas D.	Analysis of Small Caliber Manueverable Projectile (SCMP) Concepts for Helicopter and Air Defense Applications	2	141
Huxsoll, David L.	See Twartz, John C.	3	411
Hynes, John N. Jimarez, David S.	Heuristic Information Processing as Implemented in Target Motion Resolution Analysis of Radar Data	2	157
Hynes, Thomas V.	Attenuation of High Intensity Reradiated Light by Photochromic Glass	2	171
Iafrate, Gerald J.	Utilization of Quantum Distribution Functions for Ultra-Submicron Device Transport	2	177
Jamison, Keith A.	See Thomson, George M.	3	385
Jamison, Robert G.	See Chapin, Charles C.	1	209
Jenkins, Thomas	See Lunardini, Virgil J.	2	263
Jenkinson, Howard A. Zavada, John M.	CO_2 Laser Waveguiding in GaAs MBE Layers	2	189
Jimarez, David S.	See Hynes, John N.	2	157
Johnson, John L. Herren, Kenneth A. Morgan, Robert L. Tanton, George A.	Active Imaging of Range Targets at 1.2 Millimeters	2	199
Johnson, Robert A.	See Schwering, Felix K.	3	201
Kane, P. J.	See Singler, R. E.	3	297
Kapsalis, John G.	See Porter, William L.	3	31
Kerstein, Morris	See Hubbard, Roger W.	2	125
Kirkwood, B.	See Figucia, F.	1	383
Kittleson, John K. Yu, Yung H.	Holographic Interferometry Technique for Rotary Wing Aerodynamics and Noise	2	209

Author	Title	Vol	Page
Kiwan, Abdul R.	See Howe, Philip M.	2	109
Klinger, Liliana Griffith, James R.	Fluoropolymer Barriers to Stress Corrosion in Optical Fibers	2	223
Knudson, Gregory B.	See Mikesell, Perry	2	385
Koulouris, T. N.	See Singler, R. E.	3	297
Koza, W.	See Figucia, F.	1	383
Kronenberg, Stanley	Tactical Gamma and Fast Neutron Dosimetry with Leuko Dye Optical Waveguides	2	235
Krouse, Michael R.	See Goicoechea, Ambrose	1	475
Krouse, Michael R.	See Goicoechea, Ambrose	1	489
Kunkel, Kenneth E.	See Walters, Donald L.	4	239
LeDuc, James	See Lemon, Stanley M.	2	249
Lee, H.	See Singler, R. E.	3	297
Lemon, Stanley M. LeDuc, James Binn, Leonard N.	Isolation of Hepatitis A Virus from *the New World Owl Monkey:* A New Animal Model for Hepatitis A Infections	2	249
Lewis, Danny H.	See Setterstrom, Jean A.	3	215
Lewis, James H.	See Ashman, William P.	1	1
Lieberman, Michael M.	See Powanda, Michael C.	3	63
Link, Lewis E.	See Collins, John G.	4	47
Link, Lewis E.	See Gladen, Curtis L.	1	459
Liu, H. K.	See Duthie, Joseph G.	1	341
Lunardini, Virgil J. Berg, Richard McGaw, Richard Jenkins, Thomas Nakano, Yoshisuke Oliphant, Joseph O'Neill, Kevin Tice, Allan	The Mobility of Water in Frozen Soils	2	263

Author	Title	Vol	Page
Lund, David J. Beatrice, Edwin S. Schuschereba, Steven	Bioeffects Data Concerning the Safe Use of GaAs Laser Training Devices	2	279
Lund, David J.	See Stuck, Bruce E.	3	371
Lund, David J.	See Zwick, h.	3	449
Machuca, Raul	Computer Detection of Low Contrast Targets	2	293
Mager, Milton	See Hubbard, Roger W.	2	125
Malik, Roger J. AuCoin, Thomas R. Ross, Raymond L. Savage, Robert O.	The Planar Doped Barrier: A New Class of Electronic Devices	2	309
Mando, Michael A. Heise, Carl J.	Development of Armature Insulation Technique for Compact, High Power Alternators	4	111
Manriquez, R.	See Bushell, M.	1	159
Marchese, Vincent P.	IR Algorithm Development for Fire and Forget Projectiles	2	325
Marchionda, Kristine M.	See Voorhees, James W.	3	425
Marcucci, P. J.	See Ayers, O. E.	1	17
Martel, C. James	Development of a New Design Procedure for Overland Flow System	2	341
McCown, J. M.	See Henchal, Erik A.	2	61
McCreery, M. J. Swenberg, Charles E. Basso, Michael J. Conklin, James J. Hsieh, Jen-Shu	Biologic Dosimetry for Nuclear Environments by Electron Paramagnetic Resonance (EPR) Methods	2	357
McDonough, John H.	The Effects of Nerve Agents on Behavioral Performance and Their Modification with Antidotes and Antidote Combinations	2	371
McGaw, Richard	See Lunardini, Virgil J.	2	263

Author	Title	Vol	Page
McKnight, W. B.	See Barr, Thomas A., Jr.	1	47
Meliza, Larry L.	See Hiller, Jack H.	2	75
Melvin, W. S.	See Ayers, O. E.	1	17
Merkel, G.	See Bushell, M.	1	159
Meyers, William E.	See Setterstrom, Jean A.	3	215
Mikesell, Perry Knudson, Gregory B.	Plasmids of Legionella Species	2	385
Miller, Miles C.	Flight Instabilities of Spinning Projectiles Having Non-Rigid Payloads	2	393
Mirabelle, Rosemary	See Groff, John N.	4	97
Morgan, Robert L.	See Johnson, John L.	2	199
Murfree, J. A.	See Ayers, O. E.	1	17
Murphy, Newell R.	Armored Combat Vehicle Technology (ACVT) Program Mobility/Agility Findings	2	409
Nakano, Yoshisuke	See Lunardini, Virgil J.	2	263
Nietubicz, Charles J. Sturek, Walter B. Heavey, Karen R.	Computations of Projectile Magnus Effect at Transonic Velocities	2	425
Nolan, Raymond V.	Explosives Detection Systems Employing Behaviorally Modified Rats as Sensory Elements	2	441
Nomiyama, N. T.	See Rohde, R. S.	4	177
Norton, M. C.	See Rohde, R. S.	4	177
Obert, Launne P.	See Ratches, James A.	4	163
O'Connell, Robert L.	Soviet Land Arms Acquisition Model	4	125
O'Neill, Kevin	See Lunardini, Virgil J.	2	263
Oliphant, Joseph	See Lunardini, Virgil J.	2	263

Author	Title	Vol	Page
Patton, John F. Vogel, James A. Bedynek, Julius Alexander, Donald Albright, Ronald	Aerobic Power and Coronary Risk Factors in 40 and Over Aged Military Personnel	3	1
Pearson, Richard J.	See Gregory, Frederick H.	1	531
Pena, Ricardo	See Cohn, Stephen L.	1	239
Phelps, Ruth H.	Expert's Use of Information: Is It Biased?	3	17
Porter, William L. Kapsalis, John G. Wetherby, Ann Marie Drolet, Anne M. Black, Edward D.	A Rationale for Evaluation and Selection of Antioxidants for Protection of Ration Items of Different Types	3	31
Poston, Alan M. Garrett, Paul F., Jr. DeBellis, William B. Reed, Harry J. Garinther, John M.	Human Engineering Laboratory Aviation Supply Class III/V Materiel (HELAVS III/V) Field Test	3	47
Powanda, Michael C. DuBois, John Villarreal, Ysidro Lieberman, Michael M. Pruitt, Basil A., Jr.	Biochemical Indicators of Infection and Inflammation in Burn Injury	3	63
Poziomek, Edward J.	See Ashman, William P.	1	1
Presles, Henri-Noel	See Harris, Paul	2	33
Prichard, Dorothy A.	See Wolfe, Alan D.	3	435
Prifti, Joseph J. DeLuca, Eugenio	Development of Ballistic Spall-Suppression Liner for M113 Armored Personnel Carrier	4	139
Pruitt, Basil A., Jr.	See Powanda, Michael C.	3	63
Ramsley, Alvin O. Yeomans, Walter G.	Psychophysics of Modern Camouflage	3	79
Randers-Pehrson, Glenn	Nonaxisymmetric Anti-Armor Warheads	4	153

Author	Title	Vol	Page
Randolph, David I. Schmeisser, Elmar T. Beatrice, Edwin S.	Laser Flash Effects: A Non-Visual Phenomenon?	3	95
Randolph, Thomas C.	See Rickett, Daniel L.	3	117
Ratches, James A. Obert, Luanne P.	FLIR/MMW Radar vs FLIR Alone	4	163
Reed, Harry J.	See Poston, Alan M.	3	47
Reed, Lockwood W.	Voice Interactive Systems Technology Avionics (VISTA) Program	3	107
Rickett, Daniel L. Adams, Nelson L. Gall, Kenneth J. Randolph, Thomas C. Rybczynski, Siegfried	Differentiation of Peripheral and Central Actions of Soman-Produced Respiratory Arrest	3	117
Rohani, Behzad	Probabilistic Solution for One-Dimensional Plane Wave Propagation in Homogeneous Bilinear Hysteretic Materials	3	131
Rohde, R. S. Buser, R. G. Norton, M. C. Dixon, R. E. Nomiyama, N. T. Chandra, S.	Laser Technology for Identification on the Modern Battlefield	4	177
Rokkos, Nikolaus	See Schwering, Felix K.	3	201
Ross, Raymond L.	See Malik, Roger J.	2	309
Roth, John A.	Measured Effects of Tactical Smoke and Dust on Performance of a High Resolution Infrared Imaging System	3	147
Rybczynski, Siegfried	See Rickett, Daniel L.	3	117
Salomon, Mark	Properties of $SOCl_2$ Electrolyte Solutions	3	163
Sandus, Oscar	See Gilbert, Everett E.	4	87
Saunders, J. Peter	See Twartz, John C.	3	411

Author	Title	Vol	Page
Savage, Robert O.	See Malik, Roger J.	2	309
Scharf, W. D.	See Bushell, M.	1	159
Schmeisser, Elmar T.	See Randolph, David I.	3	95
Schuschereba, Stephen Beatrice, Edwin S.	Autoradiography of Primate Retina After Q-Switched Ruby Laser Radiation	3	173
Schuschereba, Stephen	See Lund, David J.	2	279
Schwalm, Robert C.	See Crumley, Lloyd M.	1	283
Schwartz, Paul M.	Methods for Evaluating Gun-Pointing Angle Errors and Miss Distance Parameters for an Air Defense Gun System	3	189
Schwering, Felix K. Johnson, Robert A. Rokkos, Nikolaus Whitman, Gerald M. Violette, Edmond J. Espeland, Richard H.	Effects of Vegetation and Battlefield Obscurants on Point-to-Point Transmission in the Lower Millimeter Wave Region (30-60 GHz)	3	201
Selvaraju, G.	See Twartz, John C.	3	411
Setterstrom, Jean A. Tice, Thomas R. Lewis, Danny H. Meyers, William E.	Controlled Release of Antibiotics from Biodegradable Microcapsules for Wound Infection Control	3	215
Share, Stewart	See Baba, Anthony J.	4	15
Sharp, Edward J. Garn, Lynn E.	Electrical Properties of Heated Dielectrics	3	227
Sheldon, William J.	The Development and Production of the Tungsten Alloy M74 Grenade	4	193
Shirai, Akira	See Twartz, John C.	3	411
Shuely, Wendel J.	A New Interactive, Computer-Controlled Method for Investigating Thermal Reactions for the 'Thermodetoxification' and Resource Recovery of Surplus Chemicals	3	243

Author	Title	Vol	Page
Silverstein, Joseph D.	Near Millimeter Wave Radiation from a Gyromonotron	3	259
Simutis, Zita M. Barsam, Helena F.	Terrain Visualization by Soldiers	3	273
Sindoni, Orazio I.	Calculation on Optical Effect of Matter from First Principles Using Group Theoretical Techniques	3	283
Singler, R. E. Koulouris, T. N. Deome, A. J. Lee, H. Dunn, D. A. Kane, P. J. Bieberich, M. J.	Synthesis and Evaluation of Phosphazene Fire Resistant Fluids	3	297
Slagg, Norman	See Gilbert, Everett E.	4	87
Smith, Alvin	A Method of Polymer Design and Synthesis for Selective Infrared Energy Absorption	3	309
Sollott, Gilbert P.	See Gilbert, Everett E.	4	87
Spellicy, Robert L.	See Watkins, Wendell R.	4	249
Spoonamore, Janet H.	CAEADS--Computer Aided Engineering and Architectural Design System	3	325
Squire, Walter H. Garnett, Lamont W.	Visco-Elastic Behavior of Incendiary Compositions Under Ballistic Loading	3	341
Steiner, James	See Chu, Shih C.	1	225
Sterling, Bruce S.	The Relationship Between Company Leadership Climate and Objective Measures of Personnel Readiness	3	357
Stern, Richard A.	See Borowick, John	1	113
Stuck, Bruce E. Lund, David J. Beatrice, Edwin S.	Ocular Flash Effects of Relatively "Eye Safe" Lasers	3	371
Sturdivan, Larry M. Bexon, Roy Berkhimer, Karl	General Bullet Incapacitation and Design Model	4	209

Author	Title	Vol	Page
Sturek, Walter B.	See Nietubicz, Charles J.	2	425
Swenberg, Charles E.	See McCreery, M. J.	2	357
Tanton, George A.	See Johnson, John L.	2	199
Thomson, George M. Jamison, Keith A.	In-Bore Propellant Media Density Measurements by Characteristic X-Ray Radiography	3	385
Throop, Joseph F.	A Fracture and Ballistic Penetration Resistant Laminate	3	397
Tice, Allan	See Lunardini, Virgil J.	2	263
Tice, Thomas R.	See Setterstrom, Jean A.	3	215
Turetsky, Abraham L.	Advances in Multispectral Screening	4	225
Twartz, John C. Shirai, Akira Selvaraju, G. Saunders, J. Peter Huxsoll, David L. Groves, Michael G.	Doxycycline Prophylaxis of Scrub Typhus	3	411
Upatnieks, Juris	See Duthie, Joseph G.	1	341
Verdier, Jeff S.	See Wolfe, Alan D.	3	435
Viechnicki, Dennis J.	See Caslavsky, Jaroslav L.	1	201
Villarreal, Ysidro	See Powanda, Michael C.	3	63
Violette, Edmond J.	See Schwering, Felix K.	3	201
Vogel, James A.	See Patton, John F.	3	1
Voorhees, James W. Marchionda, Kristine Atchison, Valerie L.	Speech Command Auditory Display System (SCADS)	3	425
Wade, William L., Jr.	See Gilman, Sol	1	447
Walbert, James N.	See Elder, Alexander S.	1	353
Walters, Donald L. Kunkel, Kenneth E. Hoidale, Glenn B.	Optical Turbulence within the Convective Boundary Layer	4	239

Author	Title	Vol	Page
Watkins, Wendell R. White, Kenneth O. Spellicy, Robert L.	Simulated Plume Radiative Transfer Measurements	4	249
Wetherby, Anne Marie	See Porter, William L.	3	31
Wharton, W. W.	See Ayers, O. E.	1	17
White, Kenneth O.	See Watkins, Wendell R.	4	249
Whitman, Gerald M.	See Schwering, Felix K.	3	201
Williams, C.	See Figucia, F.	1	383
Wolfe, Alan D. Emery, Clarence E. Verdier, Jeff S. Prichard, Dorothy A.	Studies on Butyrylcholinesterase Inhibitors	3	435
Yeomans, Walter G.	See Ramsley, Alvin O.	3	79
Yergey, Alfred L.	See Friedman, Melvin H.	1	415
Yu, Yung H.	See Kittleson, John K.	2	209
Zavada, John M.	See Jenkinson, Howard A.	2	189
Zegna, Angelo I.	See DeVenuto, Frank	1	299
Zegna, Angelo I.	See DeVenuto, Frank	1	315
Zimmerman, Kathleen L.	See Elder, Alexander S.	1	353
Zwick, H. Bloom, Kenneth R. Lund, David J. Beatrice, Edwin S.	Laser Ocular Flash Effects	3	449

A DECISION TREE FOR CHEMICAL DETECTION APPLICATION (U)

*WILLIAM P. ASHMAN, Mr.
JAMES H. LEWIS, Mr.
EDWARD J. POZIOMEK, PhD.
CHEMICAL SYSTEMS LABORATORY, USAARRADCOM
ABERDEEN PROVING GROUND, MD 21010

I. Introduction

In chemical detection research, there is considerable interest in finding reactions and reagents that will be active in solid state in detecting chemical agents and agent simulants at low concentrations. The application of the reagents and their interactions would be for use in detection devices such as detector tubes, personnel dosimeters, or solid state coatings for various types of microsensor devices such as piezoelastic crystals.

In this research, a major problem is the development of a reagent and its interaction that is specific for a chemical agent or simulant without interference from other chemicals in the environment. In solid state detection research, there has been no base of information from which to draw in designing coatings that will be specific for particular chemicals. This is true generally irrespective of the nature of the chemical to be detected. Although much may be known about the chemistry of a particular molecule in solution, this same knowledge cannot be transferred routinely in predicting solid state reactions.

In finding new solid state reagents and interactions, the classical procedure is to screen randomly many compounds until the desired detection is obtained. Not only is this screening expensive; but in most cases, it is unsuccessful. Also, using this mass screen procedure, once a suitable agent/detector reagent interaction is developed, there is no way of predicting what interferences there will be without undertaking another major screening process.

This paper describes research in the design of indandione derivatives that can be used as detector reagents and as coatings for solid-state detection purposes. A chemometrics analysis (statistical, discriminant) was performed of the results obtained with one of the reagents, 2-diphenylacetyl-1,3-indandione-1 (p-dimethylaminobenzaldazine). Chemical structural and physicochemical parameters related to the molecular association mechanism that produced fluorescence activity were identified, and a structure-fluorescence enhancement model (decision tree) for the prediction of compound/indandione reagent activity was developed. The model predicts the molecular association effects with the indandione for any compound and can be used to direct the design of solid-state detection interactions. The use of the model eliminates the mass screening of compounds and aids in the prediction of potential environmental chemical interferences. The indandione research and molecular association model developed forms an important data base for future research efforts in solid-state detection technology.

II. Materials and Methodology

A. Background

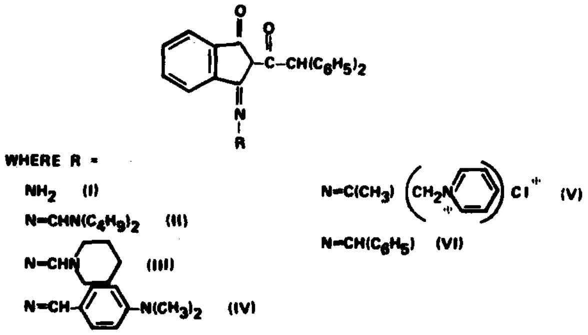

O
O
C–CH(C6H5)2
N
R
WHERE R =
NH2 (I)
N=CHN(C6H5)2 (II)
N=CHN (III)
N=CH– –N(CH3)2 (IV)
N=C(CH3) CH2N Cl⁻ (V)
N=CH(C6H5) (VI)

Figure 1. 2-DIPHENYLACETYL-1,3-INDANDIONE-1-IMINE (DIPAIN) DERIVATIVES

Crabtree and Poziomek(1-2)reported on a series of 2-diphenylacetyl-1,3-indandione-1-imine (DIPAIN) derivatives (e.g., figure 1,I-VI) that can be used in a simple, specific, and direct test (figure 2) for the detection of chemical incapacitating agents containing the α - hydroxy acid ester moiety. The detection results from a solid-state molecular association complex formed by the agent with the indandione reagent that enhances fluorescence activity. In simple terms, this means that when the DIPAIN derivative is brought into contact with certain molecules, a brilliant fluorescence occurs that is easily detectable.

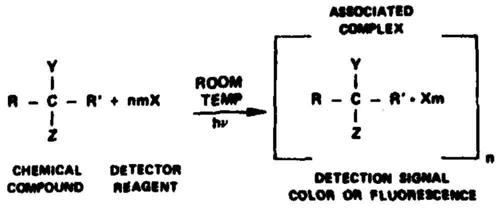

ASSOCIATED
COMPLEX
Y
R – C – R' + nmX
Z
ROOM
TEMP
hν
Y
R – C – R' · Xm
Z
n
CHEMICAL
COMPOUND
DETECTOR
REAGENT
DETECTION SIGNAL
COLOR OR FLUORESCENCE

Figure 2. CHEMICAL COMPOUND/DETECTOR REAGENT MOLECULAR ASSOCIATION EFFECT THAT PRODUCES FLUORESCENCE.

Screening studies for compound detection with 6 of the DIPAIN reagents indicated that each reagent has a distinctive fluorescence activity profile. For the compounds tested, a reagent may give a sensitive fluorescence (a strong enhancement) with some but only a moderate to no response with others (figure 3).

	FLUORESCENCE INTENSITY[a]						
COMPOUND	CMPD ALONE	I	II	III	IV	V	VI
INSECTICIDES							
DDT	NEG	M	S	MS	S	NEG	M
CHLORDANE	M	M[b]	M	MS	S[c]	W[d]	M
HEPTACHLOR	M	W	M	M	S[c]	W	M
TOXAPHENE	M	W	M	M	S[b]	W	M
RODENTICIDE							
WARFARIN	W	M	S[b]	M	MS[c]	M[b]	M
HERBICIDE							
PHENOXYACETIC ACID	M	NEG	MS	S[b]	W[c]	MS	M

[a] OBSERVED RELATIVE FLUORESCENCE: NEG, NO FLUORESCENCE; W, WEAK; M, MEDIUM; MS, MODERATELY STRONG; S, STRONG. OBSERVED COLOR IN AMBIENT LIGHT: [b] YELLOW; [c] ORANGE; [d] GREEN.

Figure 3. COMPARISON OF FLUORESCENCE ENHANCEMENT EFFECTS OF DDT WITH DIPAIN DERIVATIVES.

The fluorescence enhancement detection method can be used on various surfaces; and the detection response varies with the reagent and surface used (figure 4).

	FLUORESCENCE INTENSITY					
SOLID SUPPORT	I	II	III	IV	V	VI
ALUMINA	M	M	M	S	NEG	S
CELLULOSE	S	M	W	MS	NEG	MS
GLASS FIBER	M	S	MS	S	NEG	M
WHATMAN PAPER	M	M	M	S	VW	M
SILICA GEL	W	W	–	–	–	–

OBSERVED RELATIVE FLUORESCENCE: NEG, NO FLUORESCENCE; VW, VERY WEAK; W, WEAK; M, MEDIUM; MS, MODERATELY STRONG; S, STRONG; –, NOT TESTED.

Figure 4. EFFECT OF SOLID SUPPORT ON FLUORESCENCE ENHANCEMENT WITH DDT.

For the purpose of extending the usefulness of the fluorescence enhancement technique and to obtain data that might reveal more about the mechanism of the molecular interaction responsible for the fluorescence, additional compound/indandione derivative interaction profiles of other classes of compounds (insecticides, rodenticides, amino acids, aliphatics, alcohols, nucleic acids, etc.) were determined. A comprehensive listing of the reagents, test procedures, compound/detector reagent fluorescence activity profiles, methods of synthesis, and a chemometrics analysis to develop compound/DIPAIN reagent structure fluorescence activity models have been reported.

B. Materials and Detection Assay

The solid-solid molecular association fluorescence enhancement method of detection (figure 2) is relatively simple and easy to use. It has an advantage over other chemical detection methods in that it avoids chemical reactions which usually require prolonged heating and/or concentrated acid or base to produce a detection signal. In this paper, DIPAIN IV, the p-dimethylaminobenzaldazine derivative, and its interactions are analyzed.

Approximately 750 compounds were tested with the DIPAIN IV reagent. The test procedure was as follows: liquid samples of the compounds to be detected were spotted without dilution on Gelman ITLC glass fiber sheets using 1μl disposable micropipettes. Solid samples were spotted as 10% solution (w/v) in tetrahydrofuran. Each sheet was lightly sprayed with the DIPAIN IV reagent solution (made up as a 0.1% (w/v)) solution in benzene (0.12g/l). After the solvent evaporated, the spots were scanned visually for any change in color or fluorescence. A Chromato-Vue Cabinet with XX-15C long wave (peak at 366nm) ultraviolet lamp (Ultraviolet Products, Inc., San Gabriel, CA 91778) was used to check for fluorescence. A visual observation was made, and the response was 'qualitatively' categorized into one of 8 groups (none; very weak; weak; weak medium; medium; medium strong; strong; absorbed). In the chemometrics analysis, the 8 groups were reduced into 2 categories (unacceptable fluorescence; none, very weak, weak medium, medium, absorbed; and acceptable fluorescence; medium strong, strong).

C. Chemometrics Analysis

A Chemometrics Sciences Section has been established at the Chemical Systems Laboratory (CSL) to use the theoretical chemistry and computer-aided mathematical and statistical methods as a tool to aid the chemist in the development of mechanistic models of compound activity. The following systematic approach was used:

(1) Identify by literature (3-6) and consultation with experts the compound and reagent structural and/or physicochemical features that theoretically relate to fluorescence and to the molecular association effects that may enhance or decrease fluorescence.

Two classes of features were identified: (a) Structural-Physicochemical (figure 5). These are defined by the presence or absence of a specific structural component or functional group in the compound (e.g., figure 5: X3, X12, X14) or by measuring geometric distances between specific atoms in the compound (e.g. figure 5: X6, X20).

(b) Compound/DIPAIN IV Reagent Interaction (figure 6). Depending on how the compound/DIPAIN IV complex is formed, the resultant molecular association may produce different fluorescence intensities. Figure 6

FEATURE NUMBER	FEATURE DESCRIPTOR
X1	* TYPE OF HYDROGEN BOND
X2	POTENTIAL HYDROGEN BOND
X3	CHARGE TRANSFER GROUP ($-C=C-$ GROUP PRESENT)
X4	NUMBER OF ATOMS BETWEEN HYDROGEN BOND GROUP
X5	PHENYL RING
X6	** LENGTH OF COMPOUND
X7	SUBSTITUTION ON BENZENE (I, Cl, Br, F)
X8	MULTIPLE SUBSTITUTION ON BENZENE (I, Cl, Br, F)
X9	*** N, O, S DIRECTLY ATTACHED TO BENZENE OR LARGE PLANAR RING STRUCTURE
X10	PLANAR RING SYSTEM GEOMETRICALLY > BENZENE
X11	STERIC HINDRANCE TO OH OR $\overset{O}{\overset{\|}{C}}$ GROUP OF COMPOUND
X12	$-C\equiv C-$ PRESENT
X13	$-C=C-$ OR $C=N-$ PRESENT
X14	$C-\overset{O}{\overset{\|}{C}}-C$, $-\underset{\|}{\overset{O}{\overset{\|}{P}}}-$, $-\underset{\|}{\overset{S}{\overset{\|}{P}}}-$, NOT, $\overset{O}{\overset{\|}{C}}-OH$
X15	PYRIDYL RING
X16	PHOSPHORUS PRESENT IN COMPOUND
X17	HETEROCYCLIC RING SYSTEM HAVING $C=C$
X18	$-N=N^{\pm}$
X19	TERMINAL $\overset{O}{\overset{\|}{P}}OH$, $-\overset{O}{\overset{\|}{C}}-OH$
X20	**** ABSOLUTE LENGTH OF COMPOUND

* 0 = NONE

1 = OH, NH_2, $-\overset{O}{\overset{\|}{C}}-OH$, $-\overset{O}{\overset{\|}{P}}OH$, $\overset{H}{\overset{\|}{N}}CH_3$ TERMINAL

2 = $-\overset{O}{\overset{\|}{C}}-O$, $-\underset{\|}{\overset{O}{\overset{\|}{P}}}-O$, $\overset{S}{\overset{\|}{P}}$, $\overset{O}{\overset{\|}{C}}$, $\overset{S}{\overset{\|}{P}}$, $\overset{O}{\overset{\|}{S}}$ NON-TERMINAL

** MEASURED BY COUNTING THE NUMBER OF ATOMS THAT FORM THE LONGEST CHAIN IN A 3-DIMENSIONAL STRUCTURAL MODEL OF THE COMPOUND

*** OCH_3 AND OH NOT INCLUDED

**** MEASURED BY LONGEST LENGTH IN ANGSTROM UNITS OF A 3-DIMENSIONAL DREIDING MODEL OF THE COMPOUND

Figure 5. COMPOUND STRUCTURAL/PHYSICOCHEMICAL FEATURES ANALYZED.

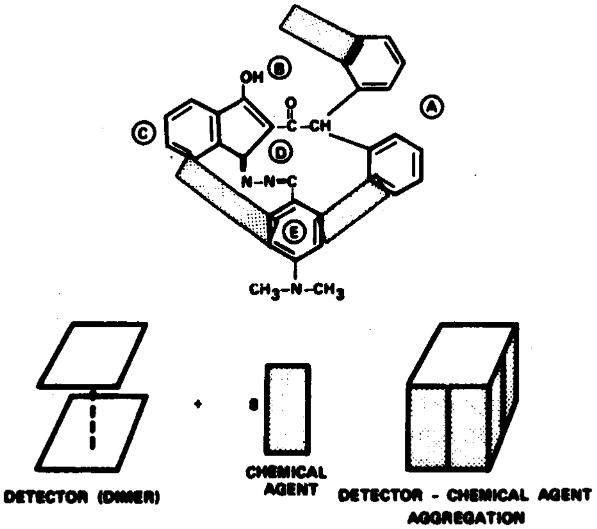
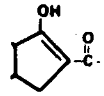


Figure 6. POTENTIAL COMPOUND/DIPAIN IV INTERACTIONS.

illustrates various ways by which a compound could interact with the DIPAIN IV reagent. For the analysis, five compound/DIPAIN IV interaction regions (Figure 6, A-E) were defined; and rules were established to determine the compound interactions with DIPAIN IV at the specific region or regions. (e.g., Region B is defined by the  area of DIPAIN IV. In order for a compound to complex with DIPAIN IV in this area, it must contain a functional group (e.g., C(O), P(O), OH, S(O) that has a potential for hydrogen bonding with the OH or -C(O) Group of Region B). Note (figure 6), a compound may interact at more than one region of the DIPAIN molecule (e.g., a compound may interact at both region A plus region E simultaneously); or, stoichiometrically, more than one compound molecule or DIPAIN molecule may complex. In the analysis, the enol form of DIPAIN IV (figure 6) is used.

(2) Select a random set of compounds from those tested with DIPAIN IV. Characterize the compounds into the defined features; mathematically depict the features and form a matrix for analysis.

206 compounds were selected for analysis. 3-dimensional Dreiding stereomodels of the structure of each compound was constructed. Each structure was used to identify whether the compound had the figure 5 features. Each feature was encoded numerically as a 1 for its presence or as a 0 for its absence. If the feature was a geometric distance or a count of the number of atoms in a specific functional relationship in the molecule, it was represented as the numerical equivalent (e.g., figure 5: X20, 7: X4,5).

In characterizing and encoding the compound/DIPAIN IV interaction features, a Dreiding model of the DIPAIN IV reagent was constructed and positioned in a fixed conformation. The DIPAIN IV molecule is flexible (not rigid) and potentially can exist in many conformations. Various conformations were constructed and analyzed. Based on the analyses, the results (section III) are optimized when the DIPAIN IV stereomodel is positioned so that the structure (figure 6) is essentially planar and there is a distance of 7.2 angstroms between the center of the phenyl ring of region E and the center of the phenyl ring of region A closest to region E. A Dreiding model of the compound is then superimposed on this DIPAIN IV model; and the interaction region or regions of DIPAIN IV that the compound may complex with are identified. In encoding, the convention of using a 1 for complex formation and a 0 for no complex is used. Seven potential compound/DIPAIN IV interactions were analyzed. These interactions were 5 complexes at regions A, B, C, D, E; and 2 compound/DIPAIN IV complexes that involved the ability of a single molecule of the compound to interact with regions A plus D simultaneously; or regions A plus E simultaneously.

A 206 compound data matrix was formed. Each vector consisted of the compound identification, the 20 features of figure 5 plus the 7 compound/DIPAIN IV interaction features, and the resultant fluorescence intensity of the compound/DIPAIN IV complex in the fluorescence enhancement empirical test.

(3) Perform computer-aided discriminant analyses of the data for the identification of the features that are related to fluorescence enhancement.

The 206 compound data matrix is divided into two sets: (a) A training set of compounds that is used to evaluate each feature and/or combination of features to establish feature discrimination values. These values are used to predict the fluorescence enhancement activity of each compound into one of the two classes; acceptable or unacceptable fluorescence. (b) A test set of compounds (different from those in the training set) that is used to determine the prediction ability of the discrimination rules established by the analysis of the training set.

The prediction ability of this discriminant analysis method is based on the number and percentage of compounds in the test set that are correctly categorized into the fluorescence classes. Various discriminant analyses using different combinations of features were performed. In each analysis, the training and test sets of compounds are structured (no two training or test sets contain an identical set of compounds) so that each analysis is unique. The BMDP7M (7) program having the option to classify using the Lachenbruch holdout or 'Jacknife Procedure' (8) is used.

(4) Develop structure-activity models for predicting the compound/DIPAIN IV fluorescence enhancement.

Using the features that were identified in the chemometrics analysis as discriminating for fluorescence enhancement, a decision tree (9-10), a partially ordered set of rules, was developed to aid in classifying the compound/DIPAIN IV fluorescence activity of untested compounds.

III. Results

A. Discriminant Analysis

Figure 7 is an example of a result that is obtained in a single computer-aided stepwise discriminant analysis. In the analysis, the parameters (features) that classify the compounds into the two fluorescence (FL) categories are identified and numerical values for each parameter to be used in future analyses are established. In the example, the training (TRNG) set is composed of 154 compounds (34 acceptable (ACC) fluorescence; 120 unacceptable (UNACC) fluorescence).

206 COMPOUNDS

TRAINING SET (154) TEST SET (52)
34 ACCEPTABLE FL 17
120 UNACCEPTABLE FL 35

CLASSIFICATION MATRIX/JACKNIFED

GROUP	% CORRECT	NUMBER OF CASES CLASSIFIED INTO GROUP	
		UNACC TRNG	ACC TRNG
UNACC TRNG	90.0	108	12
ACC TRNG	88.2	4	30
UNACC TEST	91.4	32	3
ACC TEST	94.1	1	16
TOTAL	90.3	145	61

DISCRIMINATION FEATURES:

POTENTIAL COMPOUND COMPLEX TO DIPAIN AREA (A&E)
POTENTIAL COMPOUND COMPLEX TO DIPAIN AREA (A&D)
PRESENCE OF TERMINAL $\overset{\overset{O}{\|}}{C}OH$ OR $\overset{\overset{O}{\|}}{P}OH$ GROUP

Figure 7. EXAMPLE OF DISCRIMINATION ANALYSIS RESULTS.

The discrimination analysis identifies three features as being important for categorizing the training set into the two classifications. These features are (1) the ability of a molecule of the compound to complex simultaneously with DIPAIN IV at regions A and D; (2) the ability of a molecule of the compound to complex simultaneously with DIPAIN IV at regions A and E (figure 6); and (3) the presence of a terminal C(O)OH or P(O)OH in the compound. For the training set, the discriminant analysis categorizes correctly 108 of the 120 compounds that had UNACC fluorescence and 30 of the 34 compounds that had ACC fluorescence. For the test set, 32 of the UNACC compounds and 16 of the 17 ACC compounds were correctly categorized. This is a correct percent classification rate of (48 of 52) 92.3% for the test set and a correct classification rate of (186 of 206) 90.3% for the total set.

In the final analyses, the correct classification rates for six test sets of compounds were: 90.2%, 92.3%, 91.1%, 85.7%, 86.5% and 89.6%. In all six analyses, the results indicate that the feature most important for fluorescence enhancement was the ability of a molecule of the compound to complex simultaneously with regions A and E of the DIPAIN IV reagent.

B. Decision Tree

Figure 8 is a decision tree developed from the chemometrics analysis of compound/DIPAIN IV interactions. The tree can be used to predict the solid state fluorescence (FL) enhancement of a compound with DIPAIN IV. The following decision tree rules and directions were developed.

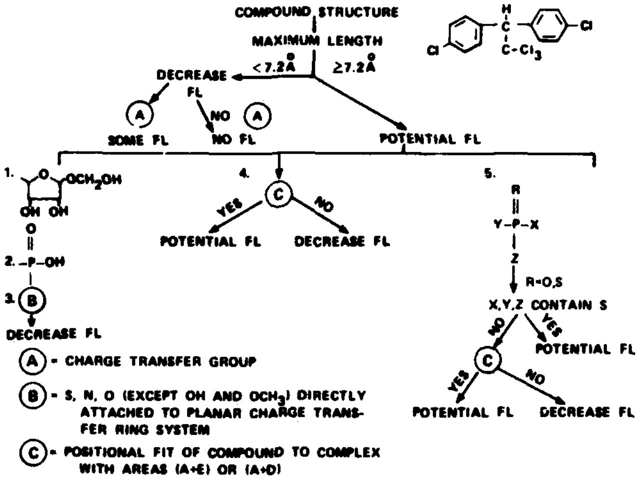

COMPOUND STRUCTURE
MAXIMUM LENGTH
<7.2Å
≥7.2Å
DECREASE FL
A
SOME FL
NO
A
NO FL
POTENTIAL FL
1.
OCH2OH
OH OH
2. –P–OH
3. B
DECREASE FL
4.
C
YES
NO
POTENTIAL FL
DECREASE FL
5.
R
Y–P–X
Z
R=O,S
X,Y,Z CONTAIN S
NO
YES
POTENTIAL FL
C
YES
NO
POTENTIAL FL
DECREASE FL
Cl
Cl
H
C
C-Cl3
A = CHARGE TRANSFER GROUP
B = S, N, O (EXCEPT OH AND OCH3) DIRECTLY ATTACHED TO PLANAR CHARGE TRANSFER RING SYSTEM
C = POSITIONAL FIT OF COMPOUND TO COMPLEX WITH AREAS (A+E) OR (A+D)

Figure 8. DECISION TREE: DIPAIN IV/FLUORESCENCE ENHANCEMENT.

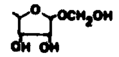
Follow these procedures and rules: (an analysis to predict the fluorescence enhancement activity of the compound DDT, illustrated in figure 8, is used as a guideline to define the procedures and rules).

Procedure (1):Construct a 3-dimensional to angstrom, Å, scale model of the structure of the compound. Measure its absolute length in angstrom units. DDT has a length of ≥ 7.2 Å.

(a) If the compounds length is <7.2Å, the compound has a high probability of not having acceptable fluorescence enhancement. Generally, if the length of the compound is <7.2 Å and if the compound has a functional group (-C=C- or -C=N-) in its structure, it may give some fluorescence enhancement with the detector reagent. However, the fluorescence enhancement intensity is predicted to be weak to medium (UNACC fluorescence). If the compound length is <7.2Å, stop. It is not necessary to proceed to procedure (2).

(b) If the compound length is ≥7.2Å, there is a higher probability of an acceptable fluorescence. DDT has a length of ≥7.2Å. Go to procedure (2).

Procedure (2): Follow steps 1.-5. as listed on figure 8. Determine if the compound has the functional groups or interaction capabilities as designated.

Step 1. The presence of the  group decreases fluorescence enhancement with DIPAIN IV. This group is highly hydrophilic and can cause steric hindrance when interacting with the DIPAIN IV. If the compound has this group, stop (UNACC FL). DDT does not contain this group. Proceed to step 2.

Step 2. The presence of P(O)OH in the compound decreases the probability of acceptable fluorescence activity. If the compound has this group, stop. DDT does not contain this group. Proceed to step 3.

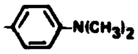
Step 3. The presence of S, N, or O (except OH and OCH3) directly attached to a planar ring system that has potential for charge transfer interaction decreases activity. If S, N, or O is attached to a large planar ring system as naphthalene with the substituent on the ring containing an S, N, or O directly attached to the ring system, the FL enhancement intensity is UNACC. The substituents OH or OCH3 do not have this result. If a compound has this configuration, stop. DDT does not have this configuration. Proceed to step 4.

Step 4.(a) If the compound contains a P(O) type configuration. Proceed to step 5. DDT does not contain Phosphorus. Proceed to 4(b).

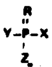
Step 4.(b) The ability of the compound to interact simultaneously with regions A plus D or regions A plus E increases the probability of ACC FL. In order to define if a compound interacts with both region A and region D simultaneously, the following rules are used:

(1) The compound must be able to bind to region A. Region A (figure 6) is defined as the diphenyl group region. In order for a compound to complex in this area, it must have a lipophilic region (at least a 3 carbon chain) or a -C=C- group in its structure. If a -C(O)O- group is present, there must be at least two carbon methylene groups attached to the -C(O)O- group, (e.g., -C-C-C(O)O-). A compound is positioned to region A by superimposing its lipophilic or -C=C group on the phenyl ring positioned closest to region D.

(2) Region D (figure 6) consists of the -C(O)-C-C=N-N=C area. A compound interacting at this region must be planar (usually a phenyl ring or a naphthalene ring configuration). After positioning the compound to region A, if the compound overlaps on to region D and contains a -C=C- group, it complexes with both region A and D.

(3) Region E consists of the  area (figure 6). The interaction rules used for region A are also used for region E. In order for a compound to interact with A plus E, the compound must be able to bind to A. After positioning it to A, it must have sufficient length to

overlap on region E. If the compound overlaps on E, an evaluation of the compound structural components that overlap on E is necessary. If the compound contains a -C=C- or -C=N- bond that can interact with E by potential charge transfer or a straight chain methylene group that can interact with E by hydrophobic type interaction, the compound is categorized as binding to regions A and E. If a compound contains a C(O)OH, P(O)OH, C(O)O-, S(O)O-, or other highly hydrophilic or potential hydrogen bonding region that overlaps on region E, it is categorized as not binding. If the compound complexes, the result is ACC FL; if it does not, the result is UNACC FL. Stop. DDT can interact with regions A plus E simultaneously. Therefore, DDT is predicted to give a strong flourescence and to be detected by a solid state detection device that uses DIPAIN IV as the detector reagent.

Step 5. This branch of the decision tree is developed for the prediction of the detection of  $Y-\overset{\overset{R}{\|}}{\underset{\underset{Z}{|}}{P}}-X$ phosphorus, P, compounds (organophosphorous nerve agents and simulants). The group R of the structure may be either a sulfur or oxygen atom. In using this part of the decision tree, orient the compound so that the substituent with the longest absolute length is positioned at X (figure 8). If X contains a sulfur atom directly attached to the P, it has a high probability of ACC FL. If it does not contain sulfur, determine if the structure can simultaneously bind to region A plus D or region A plus E. If it can interact, the result is a high probability of ACC FL. Stop.

IV. Discussion

In chemical agent detection, there is a need to develop a detection device that is specific for an agent, and that is reusable, efficient, portable, and simple to use. Solid-state microsensor devices such as piezoelastic crystals provide a technological advance that may satisfy this need. The DIPAIN reagents studied (1-2) and the compound/DIPAIN interaction fluorescence enhancement effect have application in coatings for these solid-state devices or in the production of a signal that can be used in solid-state detection.

The objectives of this study were to use chemometrics dicriminant analysis techniques to: (1) identify the mechanisms related to the fluorescence effects produced; (2) develop compound/DIPAIN IV structure-activity models that can be used as guidelines to predict the fluorescence enhancement of untested compounds; (3) aid in the design of new solid-state detection reagents that are specific for a chemical agent.

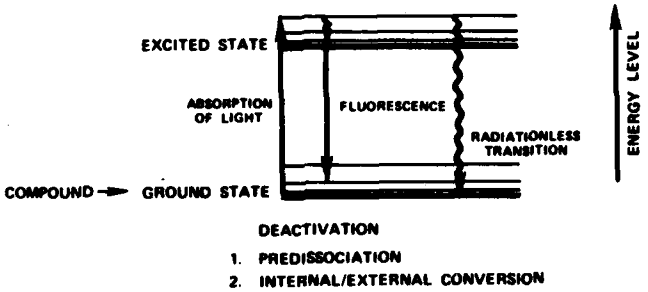


Figure 9. THEORETICAL RELATIONSHIPS FOR FLUORESCENCE ACTIVITY.

Figure 9 illustrates the excited state theory (5) and two processes associated with the deactivation of fluorescence activity:

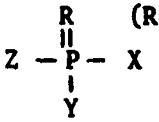
(a) Predissociation - occurs when a molecule can exist in two excited states and crossover of potential energy of the states results in a dissociation of energy;

(b) Internal/external conversion - occurs when the excess energy, provided by the ultraviolet light is converted into vibrational energy.

In the compound/DIPAIN IV interaction, the ultraviolet light excites the mixture. If predissociation and/or internal/external conversion occurs sufficiently to deactivate the complex, there is no fluorescence. The 27 features selected for analyses effect these two processes.

The DIPAIN IV molecule is flexible. In the internal/external conversion process because of this flexibility, the excess energy provided by the ultraviolet light is converted into vibrational energy; and therefore, no fluorescence results. However, if a compound can complex with the DIPAIN IV molecule with an interaction sufficient to cause rigidity of the DIPAIN IV and, thereby, prevent internal/external conversion, fluorescence results.

The results identified that the compound size (length) and the ability of the compound to interact simultaneously with the DIPAIN IV reagent at regions A plus E or A plus D influenced the fluorescence enhancement. This result suggests that the mechanism is associated with the internal/external conversion process. In a compound/DIPAIN IV interaction that produces acceptable fluorescence, the compound must have sufficient (size) length and functional groups that allows for an interaction which maintains the DIPAIN IV molecule in a rigid conformation.

The decision tree allows the user to predict the fluorescence enhancement activity with DIPAIN IV of all compounds (e.g., chemical agents, agent simulants, environmental interferences). The decision tree rules can aid in designing or identifying compounds (new agent simulants) that can be detected by DIPAIN IV. For example, in the design of a simulant for an organophosphorous agent, it is suggested to take the parent structure  $Z-\overset{\overset{\displaystyle R}{\|}}{\underset{\underset{\displaystyle Y}{|}}{P}}-X$ (R=O,S) and to use the decision tree rules to establish the position, type, and functional groups of atoms of the substituents (x, y, z) to be added to the parent configuration.

In addition to this study, the previous studies (1-2) on the DIPAIN reagents have established a data base and structure-activity relationships that can be used as guidelines in the design of new solid-state DIPAIN reagents for use in microsensor devices for the detection of a specific agent. The DIPAIN IV detection is limited to compounds that have an absolute length of $\geq$7.2Å. Many agents (e.g., GB; absolute length <7.2Å) cannot be detected by the DIPAIN IV molecule. Because the mechanism of enhancing fluorescence activity is associated with the ability of the compound to make the DIPAIN molecule rigid, it is suggested to synthesize new DIPAIN reagents having substituents that contain the basic configuration in R (figure 1) as in DIPAIN IV or VI. For this design, either basic configuration of R of IV or VI is designated RR. In order to obtain a reagent that is specific for an agent, the substituent in the RR region should be designed to allow for interaction with the functional groups or atoms of the agent molecule that do not superimpose after positioning the agent molecule to the diphenyl region of the parent DIPAIN configuration. The resultant agent/DIPAIN reagent interaction corresponds to the interaction of a compound with the DIPAIN IV molecule at regions A plus E. In the newly designed reagent, region RR is designed as a template to fit the agent.

In conclusion, the decision tree rules and compound/DIPAIN IV interaction mechanism for predicting solid state molecular association effects identified establishes a scientific basis for new concepts and potential improvements in chemical detection. The results of this study: (1) allows for the design of solid state detection interactions using DIPAIN reagents; (2) provides a set of guidelines for the development of chemical agent simulants that are detectable using the fluorescence enhancement technique; and (3) provides guidelines for the design of new DIPAIN detector reagents that are specific for a chemical agent or simulant. In summary, the DIPAIN data and results have provided fundamental information for chemical detection research which forms a basis for many potential advances in solid state detection technology.

References:

1. Poziomek, E. J., Crabtree, E. V. and Mackay, R. A. Anal. Let., 13 (A14) 1249-1254 (1980).
2. Poziomek, E. J., Crabtree, E. V. and Mullin, J. W. Anal. Let., 14 (A11), 825-831 (1981).
3. Braun, R. A. and Mosher, W. A. J. Am. Chem. Soc., 80, 3048 (1958).
4. Mosher, W. A., Bechara, I. S. and Poziomek, E. J. Talanta, 15, 482-484 (1968).
5. Joffe, W. H. and Orchin, M. Theory and Applications of Ultraviolet Spectroscopy. John Wiley & Sons, Inc. (1962).
6. Guilbault, G. G., Editor. Fluorescence, Theory, Instrumentation, and Practice. Marcel Dekker, Inc., New York. (1967).
7. BMDP-7M, Biomedical Computer Programs P-Series, Univ. of California Press. Berkeley, London, Los Angeles. Version Located on Univac 1100, Chemical Systems Laboratory, July 1981.
8. Lachenbruch, P., Biometrics, 23, 639, 1967.
9. Duda, R. O. and Hart, P.E. "Pattern Classification and Scene Analysis" John Wiley and Sons, New York, 1973.
10. Gane, C. and Sanson, T. Structured Systems Analysis: Tools and Techniques. McDonnell Douglas Corporation. 1981.

TITLE: A DECISION TREE FOR CHEMICAL DETECTION APPLICATION (U)

The use of trade names in this report does not constitute an official endorsement or approval of the use of such commercial hardware or software. This report may not be cited for purposes of advertisement.

Disclaimer

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

LASER INDUCED POLYMERIZATION REACTIONS IN SOLID PROPELLANT BINDERS (U)

O. E. Ayers, Ph.D., B. D. Allan, Ph.D., W. S. Melvin, Ph.D.,
J. A. Murfree, Ph.D., W. W. Wharton, Ph.D., and P. J. Marcucci, 2nd LT

US ARMY MISSILE COMMAND
REDSTONE ARSENAL, AL 35898

INTRODUCTION

The use of organometallic compounds to catalyze urethane reactions between polyols and isocyanates has found extensive use in many industrial and military applications. Organometallic compounds are used as catalysts in the processing of solid propellants to accelerate the curing reaction of the binder in systems where crosslinking involves urethane formation between a prepolymer with hydroxyl groups and polyisocyanates.

It has been observed that simple alcohol-isocyanate reactions are susceptible to photochemical enhancement when exposed to a low intensity tungsten light source (1). This photochemical rate enhancement of urethane formation occurred in dilute solutions of carbon tetrachloride both in the presence and absence of organometallic cure catalysts. However, the rapid polymerization of polyols with isocyanates requires the presence of an organometallic catalyst; thus, the concept of photoassisting the reactivity of these catalysts appeared to be an interesting area of investigation. Consideration of this phenomenon prompted the idea that a high energy monochromatic radiation source might be useful in promoting specific chemical reactions and better controlling the cure process of urethane reactions.

This effort examines the effect of ultraviolet and visible laser radiation in conjunction with various organometallic catalysts on the formation of propellant binders. Laser controlled reactions may proceed by pathways not encountered under thermal or other photochemical conditions. Also, the laser has the advantage of providing an intensely concentrated monochromatic beam which allows polymerization to occur at ambient temperature rather than at elevated temperatures. The potential use of the laser as a polymerization tool may provide greater control over the reaction process.

The work described here was initiated to determine the effects of laser radiation on the condensation reactions of polyols and isocyanates and to assess the influence of small concentrations of photoabsorbing additives. Several classes of organometallic compounds known for their ability to catalyze these types of reactions were screened using ultraviolet lamps. Of these, several compounds were found to be particularly receptive to photoactivation in these reactions and were subjected to detailed laser investigation. The results of monochromatic radiation on these photocatalyzed reactions at 266, 355, and 532 nm are reported here.

EXPERIMENTAL

All experimental samples were prepared by the following procedure: A sample of the organometallic compound was placed in a small mortar and ground with a pestle until the material was finely divided. One milligram of the organometallic compound was added to 4.25 ml of isophorone diisocyanate (IPDI). The sample was then thoroughly mixed by shaking for one-half hour on a mechanical shaker.

1.41 ml of the IPDI/organometallic mixture was combined with 10 ml of polyethyleneglycol adipate (PGA) and thoroughly mixed. The samples were then evacuated in a glass vacuum desiccator to remove dissolved air and then opened to the atmosphere. Some samples were run under a nitrogen or oxygen atmosphere. In these cases, the appropriate gas was used to return the samples to atmospheric pressure. An aliquot of each sample was transferred to an ultraviolet 10 mm cell for irradiation. Also an aliquot of each sample was used as a reference standard. At various time intervals, viscosity measurements were made on both the control and irradiated samples. During irradiation, the sample holder was continuously moved up and down in a vertical motion by a mechanical device to expose the entire sample to the laser beam.

All samples that were run in these experiments had an NCO/OH ratio of 1.0 and contained 0.0025 weight percent catalysts. Therefore, to maintain the same ratio, experimental samples containing hexamethylene diisocyanate (HMDI) were prepared in a similar manner to those with IPDI. Thus, only 3.3 ml of HMDI were mixed with 1.0 mg of the organometallic catalyst and then 1.07 ml of this mixture was added to 10 ml of PGA polymer to make the final sample for irradiation.

BLAK-RAY ultraviolet lamps having wavelengths of 254 and 366 nm were used in the preliminary screening of the organometallic compounds shown in Table I. Those samples which showed increased reactivity were then selected for laser testing. A Quanta-Ray Model DCR Nd:YAG laser system was used as the source of irradiation in these experiments. This DCR Nd:YAG laser provided 220 mJ/pulse at 532 nm, 115 mJ/pulse at 355 nm, and 50 mJ/pulse at 266 nm. In all the experimental reactions reported in this paper, the laser was operated at 10 pulses per second.

The Ferranti-Shirley Viscometer System was used to follow the rate of reaction of the diol prepolymer with the diisocyanate by measuring the change in viscosity as a function of time.

A Beckman Model ACTA MIV spectrophotometer was used to obtain spectra of the catalysts listed in Table II. The catalysts were used as received. Solvents used were spectroquality solvents. The solvent for ferrocene was 2,2,4-trimethylpentane. Dibutyltin diacetate solution was prepared using acetonitrile as solvent. Molar absorbtivities at 266 and 355 nm for 1,1'-ferrocenedicarboxylic acid were obtained with a methanol solution, and the molar absorbtivity at 532 nm was obtained in an acetone solution. The remaining spectra were obtained using methanol as solvent.

The preliminary screening of approximately sixty organometallic compounds indicated significant photoactivity occurred in about ten percent of the samples tested. Six samples were selected for closer examination and controlled exposure to intense monochromatic ultraviolet and visible laser irradiation. The results of these experiments are given in Tables II and V.

RESULTS

These investigations showed that ultraviolet and visible irradiation of catalyzed gumstock mixtures of polyethyleneglycol adipate (PGA) and various isocyanates could significantly increase the rates of urethane bond formation. However, the laser enhanced reaction rate was found to occur only when specific organometallic catalysts were present at the wavelengths used. This secondary acceleration in rate is over and above the already increased rate of reaction found to occur with these catalysts under nonphotolyzed conditions. At the wavelengths investigated the greatest acceleration occurred when the organometallic compounds contained iron. Most other metal containing compounds were found to produce lesser or no photorate enhancement.

Two ultraviolet lamps were employed to assist in the rapid screening of a wide variety of candidate organometallic catalysts. One lamp produced low intensity radiation at 254 nm and the other at 366 nm. These lamps permitted several samples to be simultaneously irradiated. The possible effect of photoactivity on the hydroxyl-isocyanate reactions of PGA, IPDI and various organometallic compounds was determined. After a minimum of two hours exposure, the viscosity of each sample was measured and compared with its appropriate reference sample. Those samples which showed the greatest viscosity differences were selected for further investigation with the laser. The results of these initial screening tests are shown in Table I. These viscosity differences between the irradiated and control samples were taken to be a direct measurement of the enhanced polymerization reaction.

The data presented in Table II show the actual viscosities observed, as a function of time, when identical PGA and HMDI or IPDI samples containing 0.0025% catalyst by weight were irradiated with a pulsed Nd:YAG laser. The laser enhanced reactions were observed to be extremely sensitive to the concentration of catalyst employed. When concentrations of catalyst were doubled from the 0.0025% level, the viscosity increased beyond the limit of the viscometer within 30 minutes. Solvents were not employed to avoid possible complications due to extraneous solvent-assisted charge transfer complexes which might affect the reaction mechanism.

A cursory examination was made to determine if the reactivity of a linear aliphatic isocyanate (HMDI) differed from that of a cyclic aliphatic isocyanate (IPDI) in these reactions. The data obtained do not show any reaction enhancement or inhibition trends which could be directly attributed to a photocatalyzed reaction pathway favoring one isocyanate form over the other.

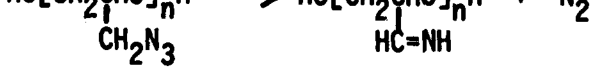
To reduce possible interferences from the presence of oxygen all samples were vacuum degassed, returned to atmospheric pressure with air, and sealed prior to irradiation. The only samples observed to undergo detectable color changes contained ferrocene as the catalyst. The faint greenish appearance was attributed to the formation of the ferrocenium ion (1). To determine if the air above the sample had an effect, samples were prepared by replacing the air with a pure nitrogen or oxygen atmosphere. Similar rates were obtained with both nitrogen and oxygen although they were slightly greater than obtained with air. As minimal effects were noted with the pure gases, air was used in conducting all other experiments.

Laser irradiation at 266 and 355 nm (Table II) was shown to have significant influence on the iron(III) acetylacetonate (FeAA) catalyzed reaction rates. To determine if this increased catalytic influence was common to other metal acetylacetonates, a series of these compounds was subjected to laser irradiation. Representative reaction viscosity data for the chromium(III), vanadyl(IV), and cobalt(III) acetylacetonates are shown in Table III and compared with the corresponding iron(III) compound. Little or no photoenhancement on the reaction rate was found with these additives and many other similar metal acetylacetonates. Table I shows additional metal acetylacetonates which were subjected to ultraviolet lamp irradiation and found to have little or no affect on these reactions.

A series of experiments was performed utilizing glycidyl azide prepolymer (GAP), which is an energetic azido containing polyol, to determine its stability in the presence of intense irradiation. It is known that the azido group absorbs strongly in the 266 nm region, and its compounds are highly susceptible to ultraviolet decomposition. Typically, alkyl azido decomposition occurs with the liberation of free nitrogen and the formation of imines as shown in reaction 1.

$$R_3CN_3 \xrightarrow{h\nu} R_2C{=}NR + N_2 \qquad (1)$$

GAP was observed to undergo ultraviolet laser decomposition in accordance with reaction 2.



$$HO\left[CH_2\underset{CH_2N_3}{CH}O\right]_n H \xrightarrow{h\nu} HO\left[CH_2\underset{HC=NH}{CH}O\right]_n H + N_2 \qquad (2)$$

Copious amounts of nitrogen gas were produced and found entrapped in the gumstock during irradiation at both 266 nm, and to a lesser extent, at 355 nm. At 532 nm, laser irradiation for up to 1.5 hours did not result in any noticable decomposition of GAP. No significant photoenhancement was observed to occur at 532 nm in the presence of FeAA.

Spectra for the six active organometallic catalysts selected for laser testing were obtained and molar extinction coefficients (Table IV) were calculated at 266, 355, and 532 nm.

DISCUSSION

The objective of these investigations was to examine the influence of monochromatic ultraviolet and visible radiation on isocyanate-polyol-organometallic catalyzed reactions to determine if reactive intermediates could be photoinduced which could affect the mechanism or rate of these reactions. Various mechanisms have been proposed for these reactions (2, 3).

The absorption of ultraviolet or visible light by a complex may cause one of three effects: (1) the rupture of a bond within the complex, (2) electron transfer between a ligand and the central metal atom, or (3) creation of an electronically excited complex. As previously mentioned, metal acetylacetonate complexes are reported to form charge transfer intermediates with isocyanates. It is proposed here that either the formation of the metal-isocyanate complex is facilitated or the complex itself is photoactivated by the laser towards attack by the OH group of the polyol. Although no detailed kinetic or mechanistic studies were undertaken during the course of this work, the results do agree with that of others in which an organometallic-isocyanate complex is postulated to be an intermediate in urethane formation.

The reactions of isocyanates with compounds containing active hydrogen are known to be profoundly influenced by catalytic additives (4). For example, it has been speculated that the reaction of an isocyanate with an alcohol in the presence of tetravalent tin catalysts occurs through the formation of an intermediate complex between the catalyst and the alcohol. The formation of a donor-acceptor bond between the oxygen of the alcohol and the tin of the catalyst activates the hydrogen on the hydroxyl group and increases the overall rate of reaction. In this situation, electrophilic attack by hydrogen on the isocyanate nitrogen could result in the formation of a strong hydrogen bond and the subsequent development of the covalent N-H bond.

Other investigators report the apparent formation of transition metal charge transfer complexes in polyol-isocyanate complexes (5). The absorption peaks are indicative of $d \longrightarrow d$ electronic transitions within a transition metal ion or of $\pi \longrightarrow \pi^*$ electronic transitions between the isocyanate and metal ion.

An examination of the summarized results in Table V give conclusive support for a laser-induced photoenhancement of metal catalyzed polyol-isocyanate reactions. In five of the six samples selected for laser examination at the three wavelengths shown, it was found that the rate enhancements were both wavelength and catalyst dependent. The sixth sample (ferrocene) showed minimal photolytic activity. It is important to emphasize that two main requirements are needed to achieve photoenhancement in these reactions. These requirements are: (a) a catalytically active ingredient, such as an organometallic compound, and (b) photoactive species. It is apparent from the data (Table V) that no photoenhancement could be induced in the non-additive containing polyol-isocyanate reaction. Also, only a limited number of the many additives which are known to exert significant chemical catalysis were found to be additionally photochemically activated at the wavelengths examined. The variations in rate can also be seen to be wavelength dependent (Table V). This wavelength dependence can be attributed to the relative ability of the complex to enter into a photoactivated or excited state that is important in the rate determining step.

In the present system, it is reasonable to assume that the mechanism is very similar to that which is operative in the non-photoactivated reaction. The basis for this assumption is that the photoenhancement effects are of secondary importance when compared to the enhanced reaction rates obtained by chemical catalyst concentration changes. Laser photoactivation of the intermediate complex provides energy necessary to further enhance the catalyzed reaction.

The observation that only a limited number of the approximately sixty metal complexes tested resulted in enhanced photochemical reactions is attributed to the lack of activation of specific metal-ligand bonds. The positive results obtained with FeAA suggested that other acetylacetonate compounds might also produce similar effects. A series of acetylacetonates were evaluated with ultraviolet lamps and, in some instances, with the Nd:YAG laser. The data presented in Tables I and III show that photochemical activation was obtained only with iron(III) acetylacetonate. However, wavelengths other than those used in these investigations may result in photoactivation of metal additives which did not show any significant rate enhancement.

An interesting observation concerning the three ferrocene derivatives listed in Table V is apparent from the percent change in viscosities of the samples. The photorate enhancement increases are in direct relation to the increasing order of the electron withdrawing ability of the groups attached to the ferrocene ring. The order is as shown.

1,1'-Ferrocenedicarboxylic Acid > Benzoylferrocene > Ferrocene

CONCLUSIONS

These studies indicate significant enhancement of the polymerization rate of organometallic catalyzed polyol-isocyanate reactions can be obtained with a Nd:YAG ultraviolet/visible laser. This enhancement is wavelength and catalyst dependent. While the precise mechanism is not known, the photorate enhancement appears to be dependent on the formation of an activated metal-isocyanate complex. With the exception of dibutyltin diacetate, the only additives which gave enhanced photochemical activity at 266, 355, and 532 nm were those organometallic compounds containing iron as the central metal atom. However, not all the iron compounds screened showed photoenhancement under the conditions of the experiments. These investigations indicate the laser can be used to enhance and control urethane formation reaction rates and opens the possibility that, through proper wavelength selection, other compounds may act as photocatalysts in these or other types of reactions.

REFERENCES

(1) S. P. McManus, H. S. Bruner, H. D. Coble, and M. Ortiz, J. Org. Chem., 42, 1428, (1977).

(2) R. S. Bruenner and A. E. Oberth, J. Org. Chem., 31, 887 (1966).

(3) D. S. Tarbell, E. C. Mallett, and J. W. Wilson, J. Am. Chem. Soc., 64, 2229 (1942).

(4) K. C. Frisch and L. P. Rumao, Rev. Macromol. Chem., 6, 103 (1971).

(5) T. E. Lipatova, L. A. Bakalo, Y. N. Niselsky, and A. L. Sirotinskaya, J. Macromol. Sci., Chem., 4, 1743 (1970).

TABLE I

ORGANOMETALLIC COMPOUNDS SCREENED BY ULTRAVIOLET (UV) LAMPS

Compounds	Viscosity (poise)		
	Reference	254nm	366nm
Hexamminecobalt(III) Chloride	56	44	63
Cobalt Phthalocyanine	50	58	60
Bis(salicylidene)ethylenediamine Chromium(III)	57	62	60
Cobaltic Acetylacetonate	43	62	53
Cobaltocene	62	69	60
Cobalt Octoate (6%)	57	69	62
Manganic Acetylacetonate	88	71	67
Manganese Carbonyl	63	60	56
Manganous Acetylacetonate Hydrate	62	56	60
Nickel Dimethylglyoxime	60	63	61
Nickelocene	55	53	50
Bis(salicylidene)ethylenediamine Nickel(II)	54	55	56
Nickel Phthalocyanine	54	55	51
Titanium Oxyacetylacetonate	60	60	62
Bis(acetylacetanato)titanium Dichloride	60	65	63
Titanium Phthalocyanine	56	52	57
Zirconocene Dichloride	56	60	61
Chromic Acetylacetonate	55	52	54
Chromium Octoate (10.5%)	54	54	55
Bis(salicylidene)ethylenediamine Chromium(III)	54	54	50
Vanadyl(IV) Acetylacetonate	100	98	85
Barium Acetylacetonate Hydrate	52	53	63
Zinc Naphthenate (8%)	52	61	54
Cerium Acetylacetonate	54	54	56
Lead Phthalocyanine	54	53	52
Potassium Nitroprusside	57	60	52
2,3-Butanedione	45	52	52
Cupric Thiocyanate	51	52	52
Cupric Hexamethylenetetramine	49	52	50
Copper Stearate	52	53	51
Copper Phenylacetylide	80	82	78
Copper Biuret	52	55	54
Copper Phthalocyanine	52	51	54
Copper Naphthenate (8%)	86	78	73
Cupric Phenylacetate	104	102	102
Cupric Ferrocyanide	61	62	61
Copper(II) Ethylacetoacetate	113	109	100

TABLE I (continued)

Compounds	Viscosity (poise) Reference	254nm	366nm
Cupric p-Toluenesulfinate	76	74	73
Copper Phenylacetylide	120	122	114
Copper 3-Phenylsalicylate	107	111	99
Cupric Dipivaloylmethane	93	106	83
Iron Naphthenate (6%)	84	82	89
Ferric Ammonium Citrate	63	62	71
*Ferric Cupferrate	78	137	140
Iron Octoate (6%)	78	83	97
Ferrous Gluconate	63	63	63
Iron Tristearate	65	63	65
*1,1'-Ferrocenedicarboxylic Acid	70	78	146
*Benzoylferrocene	88	107	122
Acetylferrocene	242	254	254
Ferrocenecarboxylic Acid	236	284	277
Bis(Hydroxymethyl)ferrocene	231	251	248
1-Ferrocenyl-1,3-butadiene	461	424	490
1,10-Phenanthroline Ferrous Sulfate (0.025M solution)	127	127	127
Iron Resinate	162	164	176
Ferric Potassium Oxalate	124	124	127
Ferric Ammonium Oxalate	138	130	133
Molybdenum Naphthenate	141	141	141
Cobalt Naphthenate	136	138	136
Cupric Naphthenate	156	159	156
*Ferrocene	98	132	122
*Dibutyltin Diacetate	217	283	221

*Selected for laser examination.

TABLE II

PGA/HMDI(IPDI) CATALYZED BY ORGANOMETALLIC COMPOUNDS

Wavelength (nm)	Time (hours)	Viscosity (poise)	
		Control	Irradiated
IRON(III) ACETYLACETONATE			
532	0.5	110	100
	1.0	178	153
	1.5	288	320
355	0.5	156	157
	1.0	299	479
	1.5	484	1021
266	0.5	122	153
	1.0	185	257
	1.5	427	808
DIBUTYLTIN DIACETATE			
532	0.5	116	104
	1.0	180	131
	1.5	261	261
355	0.5	81	81
	1.0	98	107
	1.5	123	130
266	0.5	114	115
	1.0	132	156
	1.5	275	413

TABLE II (continued)

FERROCENE			
532	0.5	61	61
	1.0	63	61
	1.5	52	61
355	0.5	66	74
	1.0	69	78
	1.5	74	88
266	0.5	60	58
	1.0	60	65
	2.0	60	81
BENZOYLFERROCENE			
532	0.5	75	81
	1.0	88	108
	1.5	95	144
	2.0	94	162
355	0.5	130	114
	1.0	-	239
	1.5	-	363
	2.0	283	438
266	0.5	46	52
	1.0	-	66
	1.5	-	89
	2.0	86	101

TABLE II (continued)

FERRIC CUPFERRATE			
532	0.5	71	87
	1.0	79	109
	1.5	81	120
	2.0	99	128
355	0.5	144	254
	1.0	-	352
	1.5	-	444
	2.0	208	583
266	0.5	55	61
	1.0	-	66
	1.5	-	84
	2.0	81	98
1,1'-FERROCENEDICARBOXYLIC ACID			
532	0.5	72	87
	1.0	78	113
	1.5	76	134
	2.0	75	166
355	0.5	52	43
	1.0	-	69
	1.5	-	92
	2.0	55	138
266	0.5	40	43
	1.0	-	43
	1.5	-	46
	2.0	55	55

TABLE III

EFFECTS OF METAL ACETYLACETONATES ON PGA/HMDI REACTION AT 355 nm

Acetylacetonate	Time (Hours)	Viscosity (poise) Control	Irradiated
Chromium(III)	0.5	54	54
	1.0	55	53
	1.5	55	56
Vanadyl(IV)	0.5	65	62
	1.0	81	76
	1.5	91	83
Cobalt(III)	0.5	55	55
	1.0	55	56
	1.5	50	62
Iron(III)	0.5	156	157
	1.0	299	479
	1.5	484	1021

TABLE IV

MOLAR EXTINCTION COEFFICIENTS

COMPOUND	266 nm	355 nm	532 nm
Ferrocene	1.48×10^3	9.50	3.17
Benzoylferrocene	5.65×10^3	6.69×10^2	1.88×10^2
1,1'-Ferrocenedicarboxylic Acid	6.82×10^3	3.53×10^2	0.0
Ferric Acetylacetonate	1.90×10^4	2.58×10^3	3.30×10^2
Ferric Cupferrate	1.85×10^4	3.67×10^3	0.0
Dibutyltin Diacetate	6.42	0.0	0.0

TABLE V

VISCOSITY CHANGES AT VARIOUS WAVELENGTHS DUE TO PHOTOCATALYTIC ACTIVITY

Catalyst	Viscosity Change in poise 266	355	532
None	0 (0%)*	0 (0%)	0 (0%)
Ferrocene	21 (35%)	133 (19%)	9 (17%)
Benzoylferrocene	14.4 (17%)	155.8 (55%)	68 (72%)
1,1'-Ferrocenedicarboxylic acid	0 (0%)	83.5 (151%)	86 (121%)
Ferric Acetylacetonate	381 (89%)	537 (111%)	32 (11%)
Ferric Cupferrate	17.3 (21%)	375 (180%)	29 (29%)
Dibutyltin Diacetate	138 (50%)	6.9 (6%)	0 (0%)

(*)Percent viscosity change of the irradiated sample compared to the control.

THE ENGINEER MODELING STUDY

DR. D. GORDON BAGBY
US ARMY CONSTRUCTION ENGINEERING RESEARCH LABORATORY
CHAMPAIGN, IL 61820

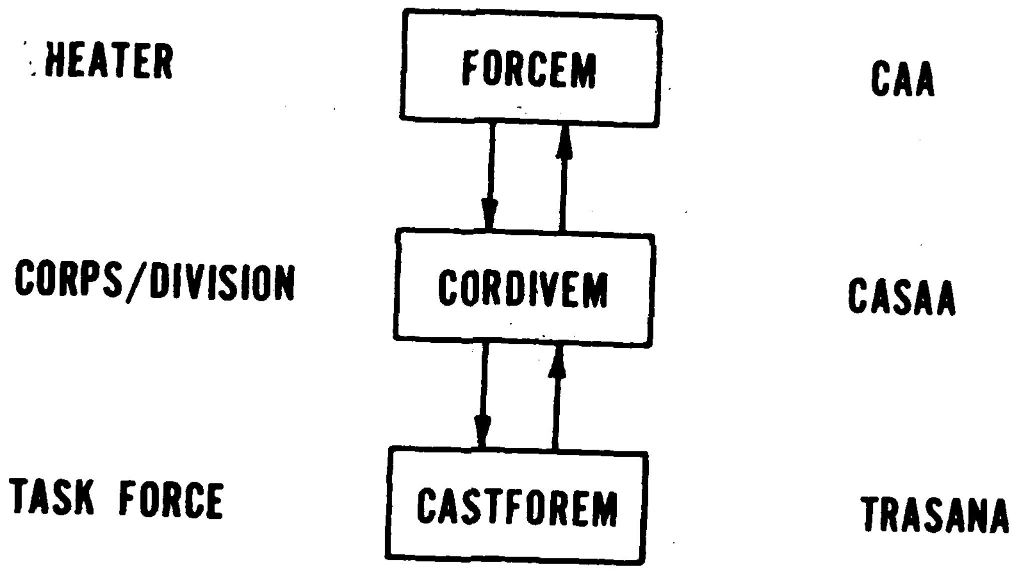
The objective of the Engineer Modeling Study is to measure the contribution of combat engineers to the effectiveness of the combined arms team. The research program originated from a Mission Area Analysis (the Engineer Family of Systems Study (E-FOSS)), which noted that Army war games were good at representing unit offensive and defensive movements, but weak in modeling the impact of US and Soviet Union combat engineer activities on battle outcomes. In 1979, the US Army Engineer School (USAES), representing the Training and Doctrine Command (TRADOC), requested the US Army Construction Engineering Research Laboratory (CERL) to correct this deficiency. The Engineer Modeling Study was the result.

The Engineer Modeling Study itself is part of the larger Army Model Improvement Program (AMIP) which seeks to improve the caliber and quality of Army war games. The three AMIP models shown in Figure 1 interact by feeding scenarios downward from higher theater-level games to lower level division and battalion-level games and by transmitting combat results upward from battalion-level games to theater-level models. The AMIP model used in the Engineer Modeling Study is the mid-level Corps/Division Evaluation Model (CORDIVEM) being developed by the Combined Arms Studies and Analysis Activity (CASAA). An implicit goal of the Engineer Modeling Study is accurate and consistent representation of the effectiveness of engineer effort throughout the AMIP model hierarchy.

Figure 2 illustrates how the Engineer Modeling Study augments the Corps/Division Evaluation Model with an Engineer Module. This module receives orders either from the CORDIVEM gamer or indirectly from within the game; the module then modifies various CORDIVEM terrain features. These terrain modifications produce three effects: mobility effects enhancing the movement of friendly troops; survivability effects reducing friendly force casualties; and countermobility effects delaying the movement of hostile forces.

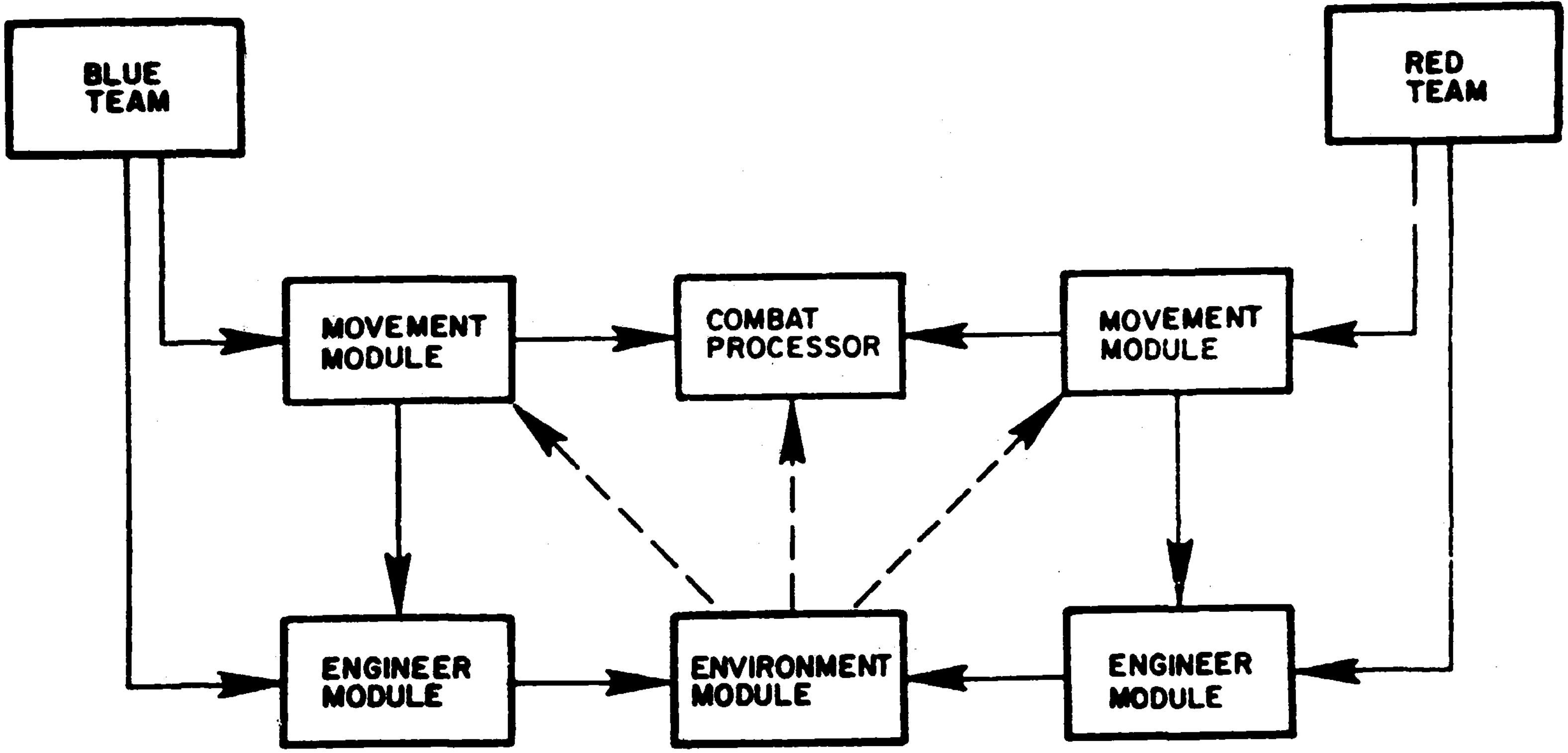
The Engineer Module is comprised of the three submodules (Figure 3):

ENGINEER MODELING STUDY (EMS) BACKGROUND


HEATER
FORCEM
CAA
CORPS/DIVISION
CORDIVEM
CASAA
TASK FORCE
CASTFOREM
TRASANA

ARMY MODEL IMPROVEMENT PROGRAM (AMIP) HIERARCHY

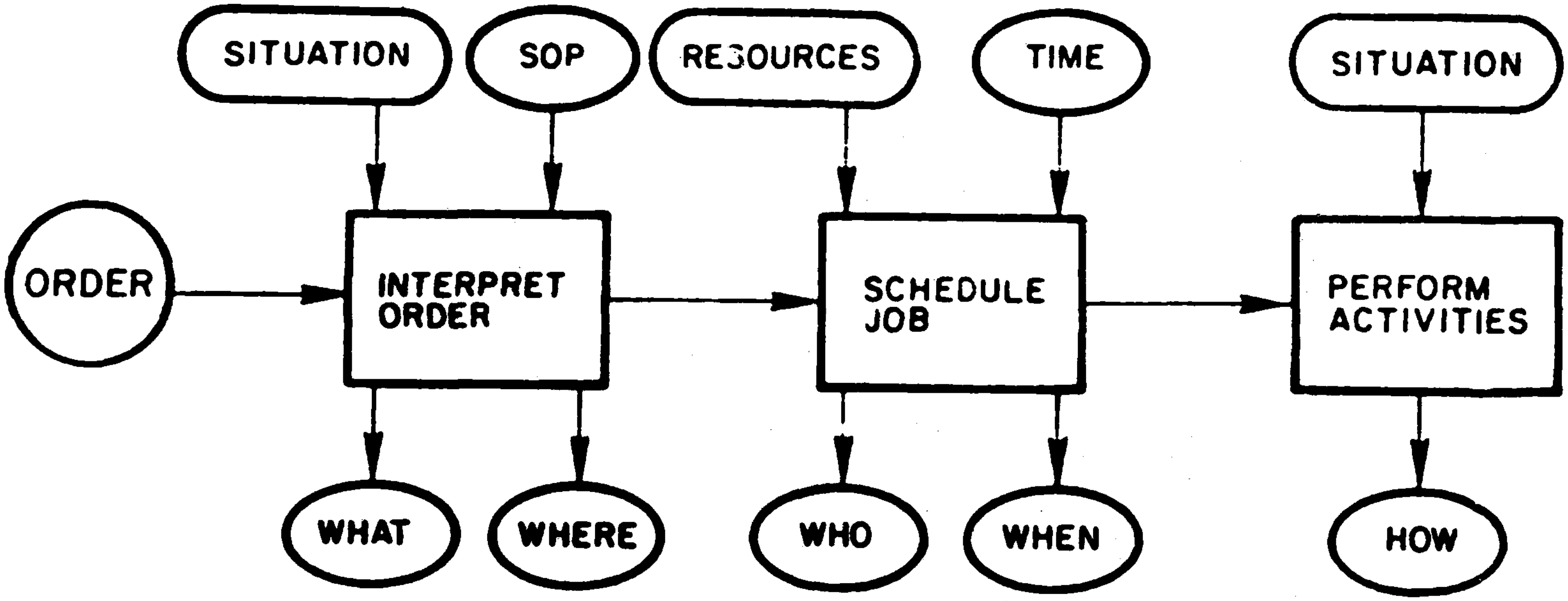
Figure 1


BLUE TEAM
RED TEAM
MOVEMENT MODULE
COMBAT PROCESSOR
MOVEMENT MODULE
ENGINEER MODULE
ENVIRONMENT MODULE
ENGINEER MODULE

CORPS BATTLE GAME STRUCTURE

Figure 2

ENGINEER PROCESS MODULE


SITUATION
SOP
RESOURCES
TIME
SITUATION
ORDER
INTERPRET ORDER
SCHEDULE JOB
PERFORM ACTIVITIES
WHAT
WHERE
WHO
WHEN
HOW

an interpret order submodule; a schedule job submodule; and a perform activity submodule. Once the Engineer Module receives orders from CORDIVEM, it interprets these orders, schedules jobs to satisfy these orders, performs the jobs, and appropriately modifies the CORDIVEM terrain.

The heart of the perform activity submodule is a set of critical path networks. For example, Figure 4 represents the precedence network for a minefield breach. Each step of this technique is shown as an activity. Given the location of the job site and the locations of the required resources, the perform activity submodule determines the completion time of the job and the resources expended.

The Engineer Module is now fully operational. It is compatible with any order stream, and can represent over 50 separate engineer activities such as blowing a bridge, building a command post, or conducting reconnaissance missions.

The schematic diagram shown in Figures 5, 6, and 7 illustrate the actual operation and interaction of CORDIVEM and the Engineer Module. Everything within the dashed lines belongs to CORDIVEM; everything outside the dashed lines is the responsibility of the Engineer Module. The time clock, or event queue, for CORDIVEM appears in the center; CORDIVEM's data files and mapboard appear on the right. The operation of the Engineer Module commences on the left with the receipt of an order from CORDIVEM, generated by either the gamer, a pregame plan, or the CORDIVEM game itself. Regardless of the source, once an order is received, it is processed by the interpret order and schedule job submodules. The interpret order submodule selects an appropriate critical path network. The schedule job submodule then places the start times for the individual events of the critical path network into the event queue of CORDIVEM's time line. The Engineer Module then goes to sleep (Figure 5). With the passage of time, CORDIVEM's TIME NOW marker moves downward. When TIME NOW reaches the start time of an activity, CORDIVEM passes control of the process back to the Engineer Module's perform activity submodule. This submodule computes a FINISH TIME and places a marker on the time line in the appropriate place (Figure 6). Finally, when TIME NOW reaches FINISH TIME, the Engineer Module's modify terrain submodule is activated and CORDIVEM's data base is changed to reflect the terrain alteration. The data base, in turn, changes the information depicted on the CORDIVEM map (Figure 7).

If, at some subsequent point in time, enemy and friendly forces engage in combat in that location, their attrition rates and movement will be modified to reflect the presence of the engineer-created terrain feature.

MI73 MINEFIELD BREACH TASK NETWORK

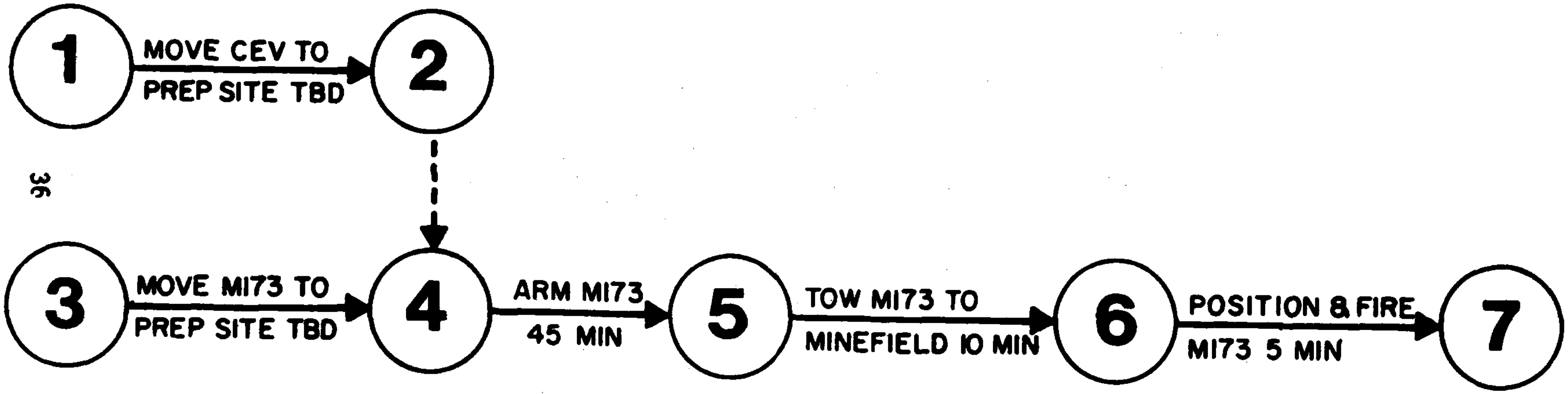

1
MOVE CEV TO
PREP SITE TBD
2
3
MOVE MI73 TO
PREP SITE TBD
4
ARM MI73
45 MIN
5
TOW MI73 TO
MINEFIELD IO MIN
6
POSITION & FIRE
MI73 5 MIN
7

Figure 4

EVENT BASED SIMULATION (STEP 1)

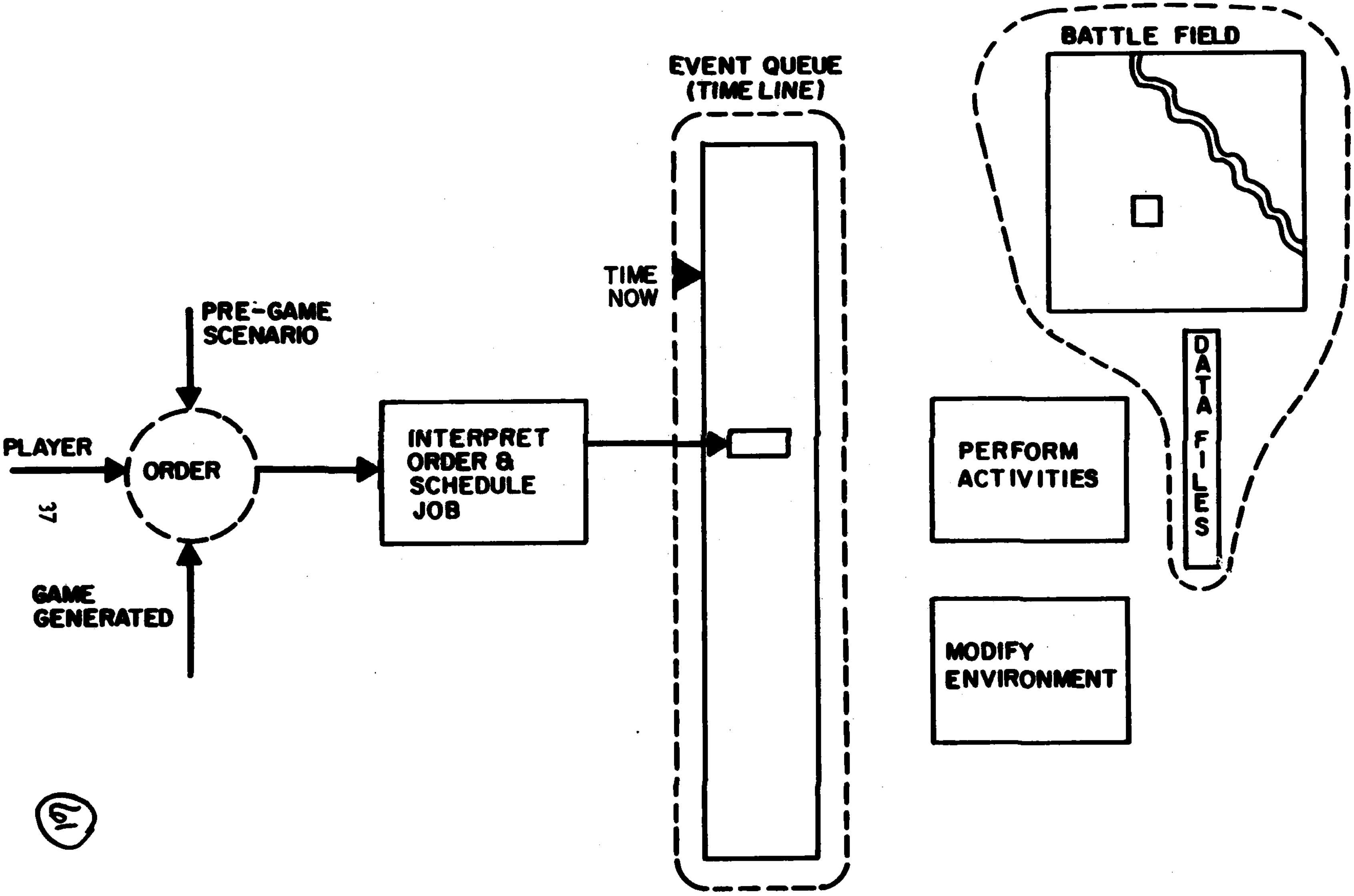

PRE-GAME
SCENARIO
PLAYER
ORDER
GAME
GENERATED
INTERPRET
ORDER &
SCHEDULE
JOB
TIME
NOW
EVENT QUEUE
(TIME LINE)
BATTLE FIELD
DATA FILES
PERFORM
ACTIVITIES
MODIFY
ENVIRONMENT

Figure 5

EVENT BASED SIMULATION (STEP 2)

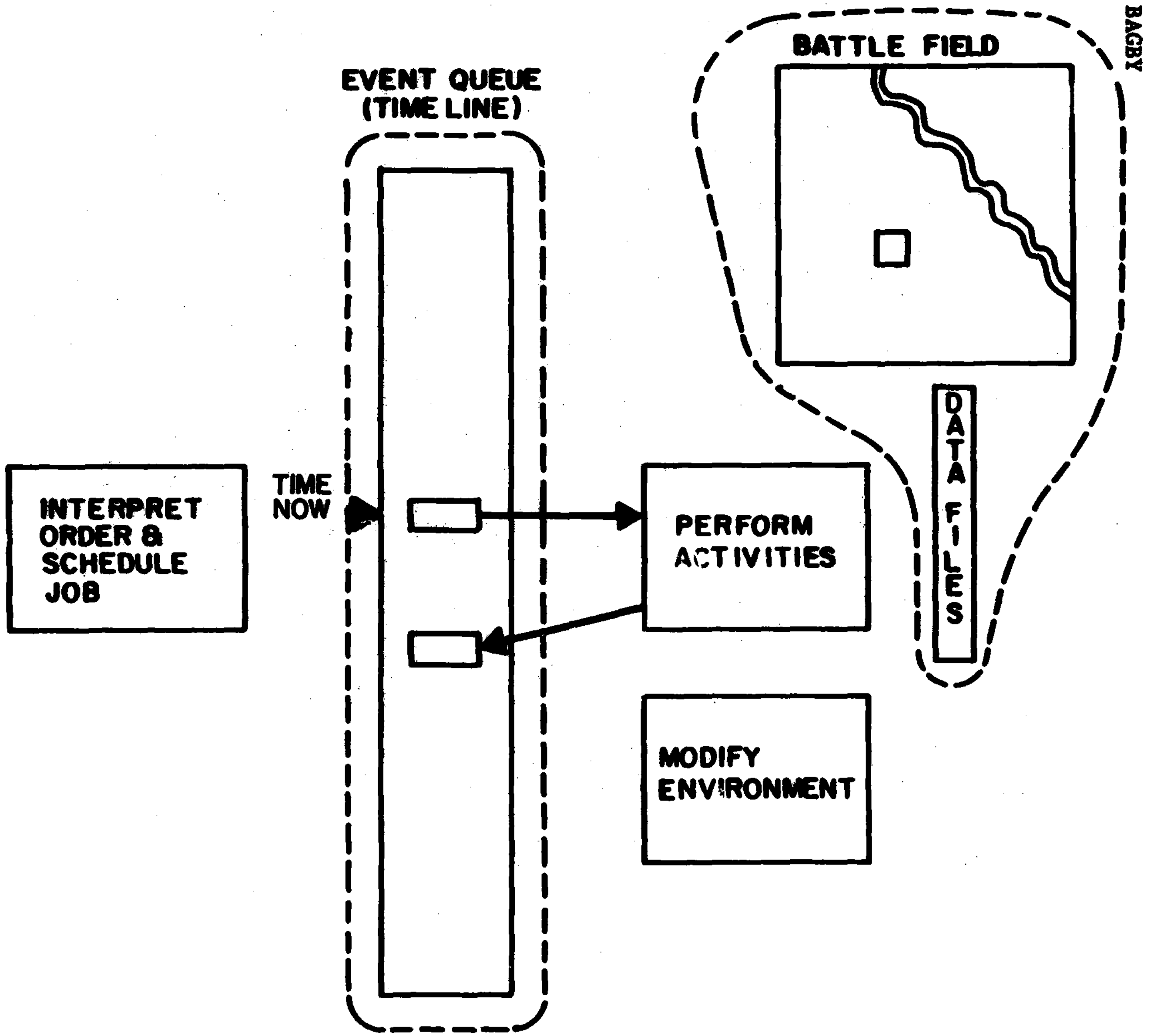

BAGBY
BATTLE FIELD
EVENT QUEUE
(TIME LINE)
DATA FILES
INTERPRET ORDER & SCHEDULE JOB
TIME NOW
PERFORM ACTIVITIES
MODIFY ENVIRONMENT

Figure 6

EVENT BASED SIMULATION (STEP 3)

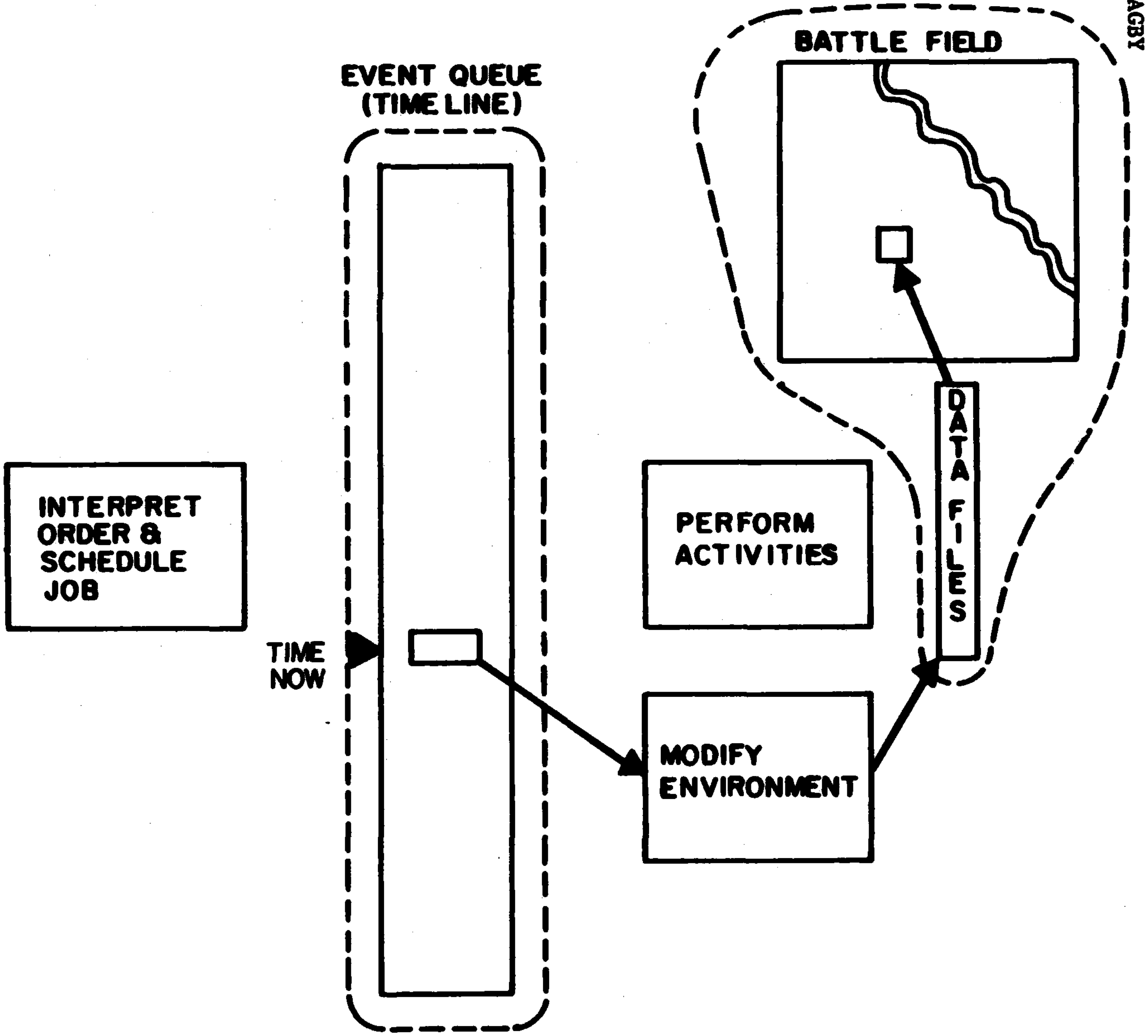

BATTLE FIELD
EVENT QUEUE
(TIME LINE)
DATA FILES
INTERPRET
ORDER &
SCHEDULE
JOB
PERFORM
ACTIVITIES
TIME
NOW
MODIFY
ENVIRONMENT

Figure 7

Figures 8, 9, and 10 are computer-drawn illustrations which depict how the CORDIVEM model ran in June 1981, both with and without combat engineer support. Figure 8 represents the initial situation at 7 AM, where five outnumbered blue armored cavalry squadrons (shown in white) confront seven red tank and reconnaissance battalions (black). By afternoon, without engineer support, blue forces were outflanked, forced to withdraw, encircled, and finally defeated by red units (Figure 9). But utilizing combat engineer support on the battlefield, bridges were destroyed and obstacles placed in the path of enemy forces to impede them until adequate defensive fortifications could be constructed (Figure 10). These efforts enabled blue forces to withstand the enemy attack for a considerably extended period, reduced blue losses, and increased red losses.

The results of these conflicts are tabulated in Figure 11. As this table indicates, when engineers were not available, red lost 73 tanks while blue suffered 62 tank casualties. With engineers present, red casualties rose to 88 tanks, while blue casualties fell to 49 tanks. It should be emphasized that these results were derived from a single scenario and many additional scenarios must be executed before the Engineer Module is credible.

During the remainder of Fiscal Year 82, sensitivity tests on the Engineer Module will be conducted and work will commence on a way to model threat engineers. Engineer Module documentation will be published containing flow charts, variable listings, definitions, information sources, and an operating manual.

In the more distant future, the Engineer Module will be used to conduct various force structure trade-off analyses and to create the following stand-alone Engineer Modules: A Force Structure Trade-Off Module for use in force design analysis; a Combat Engineer Field Module for use by combat engineer units for operational planning; and a Combat Engineer Training Module for use in training engineer officers and others.

To summarize, The Engineer Module is merely a resource allocation routine which receives orders, decrements resources, decrements time, modifies terrain, and issues statistics (Figure 12). The Engineer Module can be used to simulate the effects of different command and control structures, engineer support relationships, and engineer equipment performance characteristics. This ability to demonstrate engineer effectiveness on the battlefield should greatly assist efforts to modernize and field new engineer equipment.

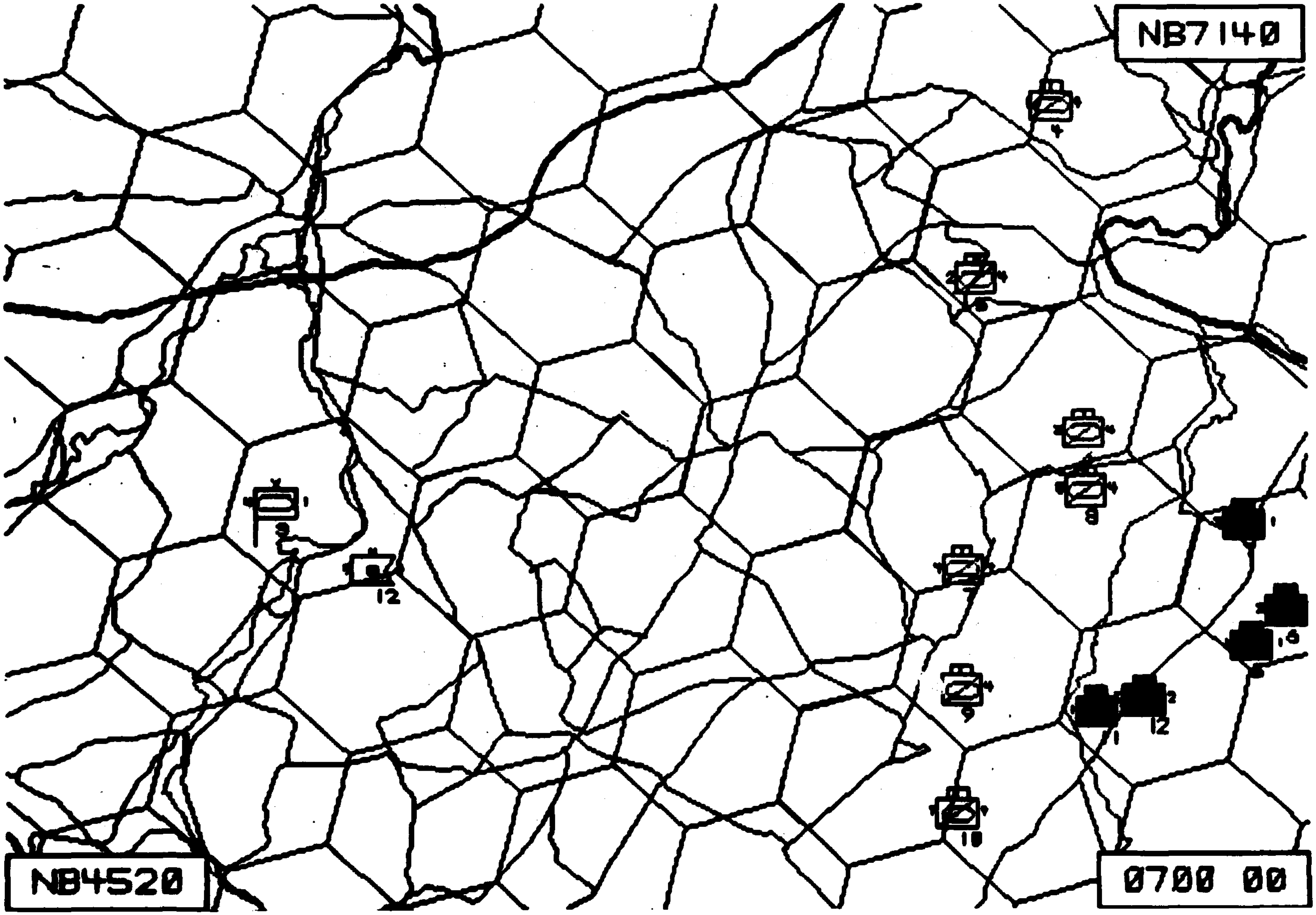

NB7140
NB4520
0700 00
12
12
1
10

Figure 8

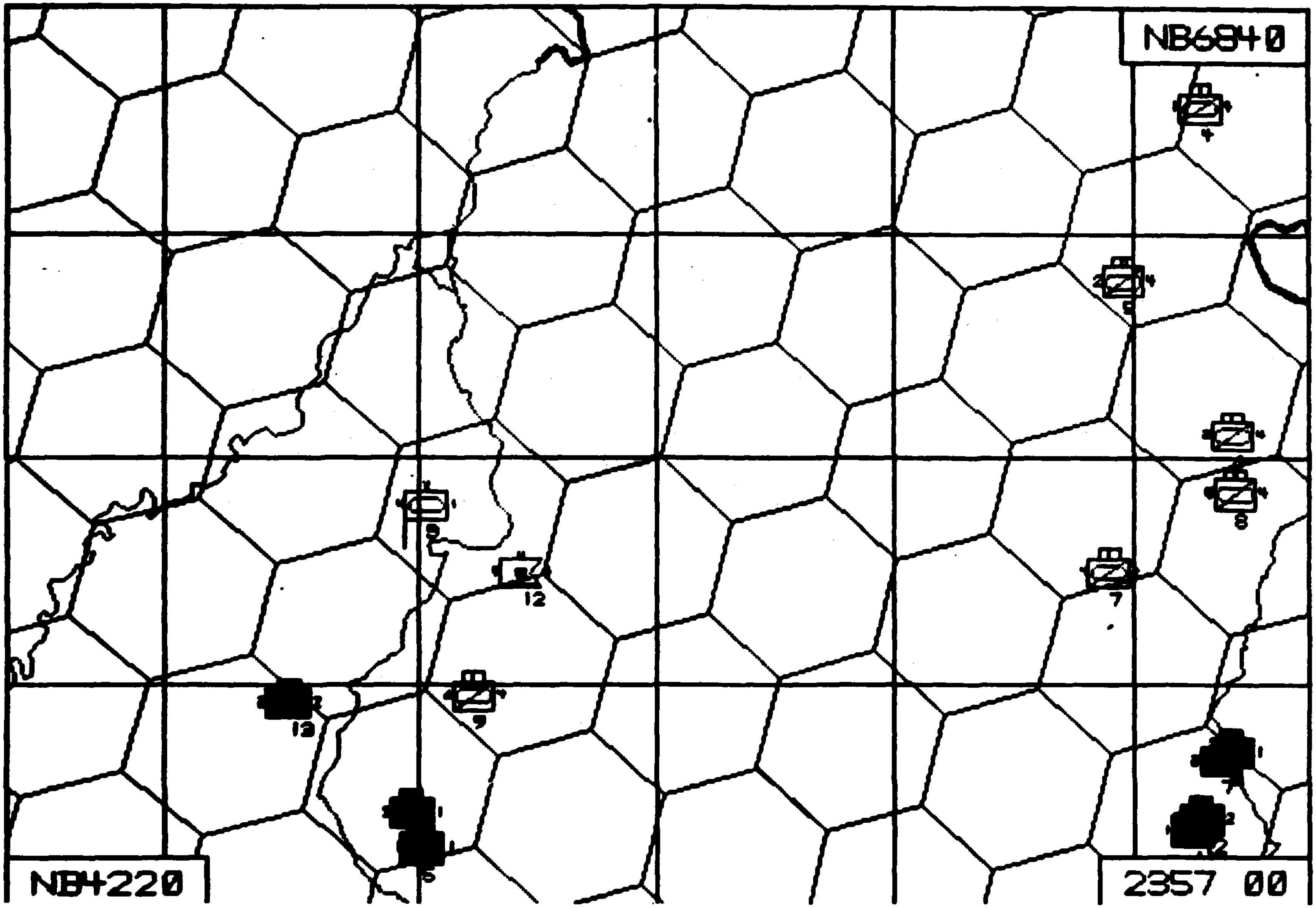

NB6840
NB4220
2357 00

Figure 9

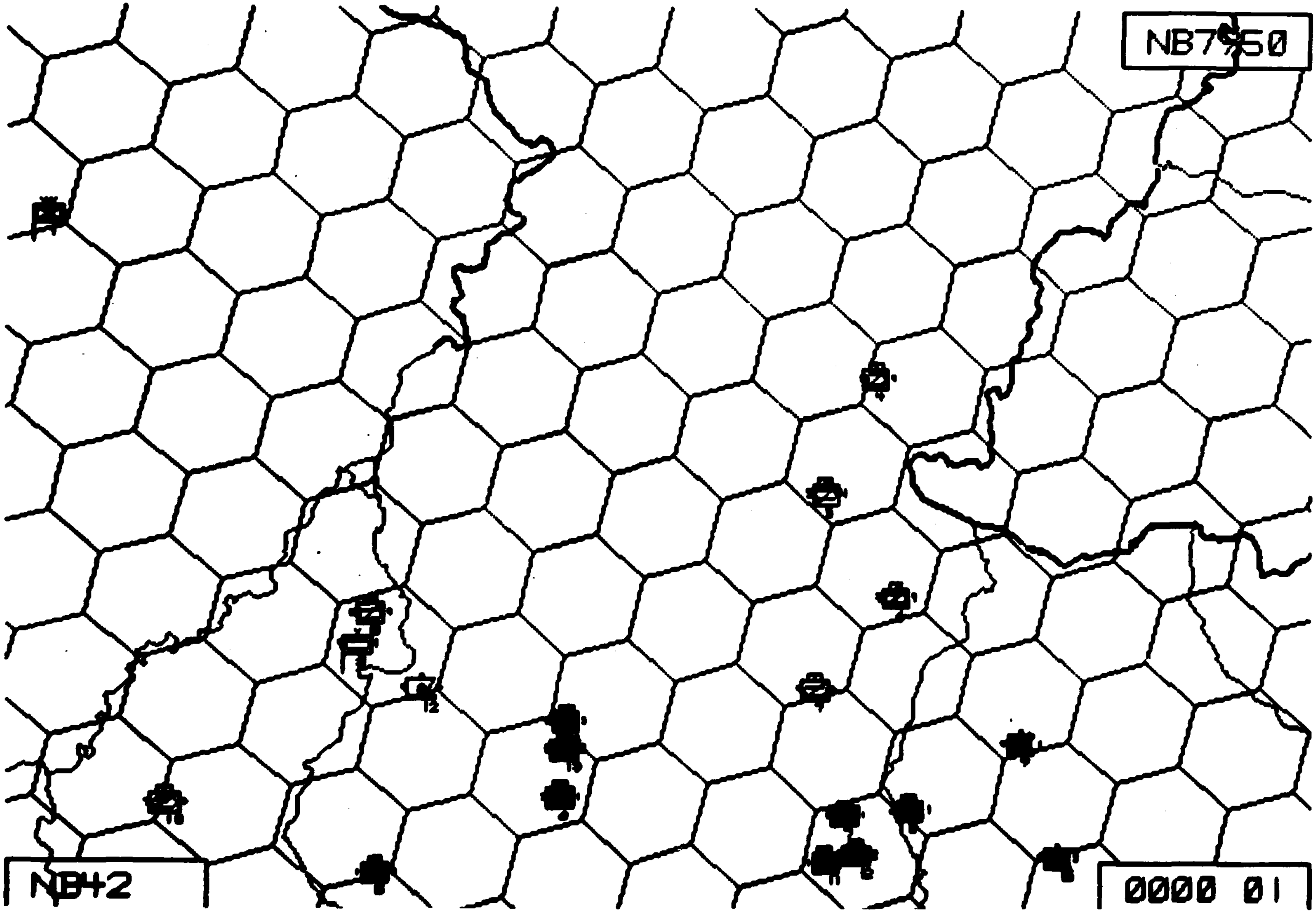

NB7950
NB+2
0000 01

Figure 10

EMS-CORDIVEM SIMULATION RESULTS

	SCENARIO WITHOUT ENGINEERS	SCENARIO WITH ENGINEERS
TIME	RED/BLUE TANK LOSSES	RED/BLUE TANK LOSSES
0800	43/31	43/31
0900	52/39	52/39
1200	55/41	53/40
1400	58/46	53/40
1600	61/49	56/41
1800	68/62	
2000	69/62	79/48
2400	71/62	86/49
END	73/62	88/49

Figure 11

ENGINEER MODEL OPERATIONS

- RECEIVES ORDERS
- DECREMENTS RESOURCES
- DECREMENTS TIME
- MODIFIES TERRAIN
- ISSUES STATISTICS

Figure 12

A NEW MIDINFRARED LASER IN HYDROGEN

THOMAS A BARR, Jr., DR.
U.S. ARMY MISSILE COMMAND
REDSTONE ARSENAL, AL 35898
W. B. McKNIGHT, DR.
UNIVERSITY OF ALABAMA IN HUNTSVILLE
HUNTSVILLE, AL 35899

Introduction: The basic research program for high energy lasers (HEL) at the U.S. Army Missile Command has had a major task in the area of new laser concepts for future Army HEL applications. A theoretical and experimental program to explore the possibility for developing a storage laser in hydrogen was carried out during FY 80-82 and resulted in the discovery of the laser reported here.

A storage laser is defined as a device which has the ability to store large amounts of energy (vibrational or electronic) without large collisional or radiative losses. There are a number of methods by which the stored energy can be extracted as stimulated emission. e.g., (i) by collisional transfer to another species with a large stimulated optical cross section, (ii) by an oscillator/amplifier cavity configuration (provided the saturation flux of the storage medium is not too large), or (iii) by temporarily increasing the stimulated cross section of the medium by such means as a field induced dipole.

Nitrogen is a prime example of a storage medium. Energy can be efficiently pumped into the vibrational mode of N_2 via discharge techniques: the collisional losses are small at room temperature and, of course, the radiative cross section for vibration/rotation transitions is zero; and nitrogen can efficiently transfer its vibrational energy to CO_2. Therefore, the N_2/CO_2 laser is an example of type (i) storage laser.

Practical examples of type (ii) storage lasers are more difficult to find because they require a value of optical cross section that is large enough to permit a realizable saturable flux and low enough to prevent rapid loss of energy by spontaneous radiation. Optical cross sections of 10^{-19}-10^{-20} cm^2 for visible transitions are required. Examples of this type of laser are the Group VI lasers - O(^{1}S), S(^{1}S), and Se(^{1}S) - which were investigated as potential drivers for fusion reactors.

The type (iii) storage lasers are, at present, only theoretical concepts. In such a laser the storage medium is pumped and, subsequently, a significant dipole moment is temporarily created in the molecule by impressing an electric field across the gas. Several investigators have proposed lasers of this type using radiation fields, as from a high intensity pulsed laser (1) or static external field from d.c. or low frequency a.c.(2). The problems in accomplishing this are formidable, and, to date, no demonstration of a field induced laser has been made.

Besides N_2, the molecules H_2 and D_2 * are good potential storage media. The available electron impact data for H_2 indicate that it should be possible to pump into the vibrational mode of H_2, D_2 and HD via an electric discharge with a controlled E/N (discharge field strength per gas density). The vibrational energy transfer kinetics of H_2 and D_2 have been experimentally measured and indicate that collisional losses at modest temperatures are small. In principle, the molecular system encompassing H_2, D_2, and HD can provide examples of all three types of storage lasers: For type (i) we have the H_2/HCl and D_2/HCl transfer electric discharge lasers (EDL), type (ii) is provided by HD itself since it has a small, permanent dipole moment, and for type (iii) H_2, D_2, and HD can all be considered for field induced dipole lasers.

Research on HD: Our work started with the analysis of HD as a type(ii), and H_2 as a type (iii) storage laser. First order calculations, in which kinetics processes were ignored, showed the conditions under which we could expect a sensible gain from either H_2 or HD based on the required partial population inversions. A major job for developing useable external fields would have been required for the H_2 laser, therefore, we elected to try the HD laser first.

Requirements for Laser Action in HD: The population distribution for threshold laser oscillation was calculated from a two level model in which the ground vibrational state, with population, Nl, was the lower state of the first vibrational state, with population, Nu, was the upper state. The required inversion density was obtained from the solution of the equations:

1. $Nu - Nl(2Ju + 1)/(2Jl + 1) = \Delta N$, and 2. $Nu + Nl = N$

Ju and Jl are the rotational quantum numbers. The required inversion density is $\Delta N = g/s$ where g is the threshold small signal gain and s is the stimulated emission cross section.

---

* D is used in this text, as commonly done, to represent deuterium.

The Einstein A coefficient was derived from data in (3); Doppler line broadening was assumed based on the physically realizable operating conditions (press = 4. torr and temp = 77°K to 300°K). Also thermal equilibrium was assumed among the rotational states. An estimate of the gain per centimeter required for threshold oscillation was made for state-of-the-art dielectric coated mirrors, R1=.995 for the back mirror and R2=.98 for the output mirror with no diffraction loss (implies a concave mirror), and a cavity length of 6.3 meters. A sensible output was assumed to exist when an absolute gain of 10^{16} was achieved, i.e., 10^{16} photons were produced in a coherent output. Threshold gain is:

3. $g = (2L)^{-1} \ln (R1 x R2)^{-1} = (1260.)^{-1} \ln (.995x.98)^{-1} = 2.0 \times 10^{-5}\ cm^{-1}$.

The stimulated emission cross sections, s, were calculated and are: for 77°K and 300°K, respectively, for P(1), $6.7x\ 0^{-22}$cm and $3.4x10^{-22}$ and for P(2) $8.9x10^{-22}cm^{2}$ and $4.6x10^{-22}cm^{2}$. Thus the lowest, and presumably easiest to get, population inversion is for the P(2) line at 77°K. In this case:

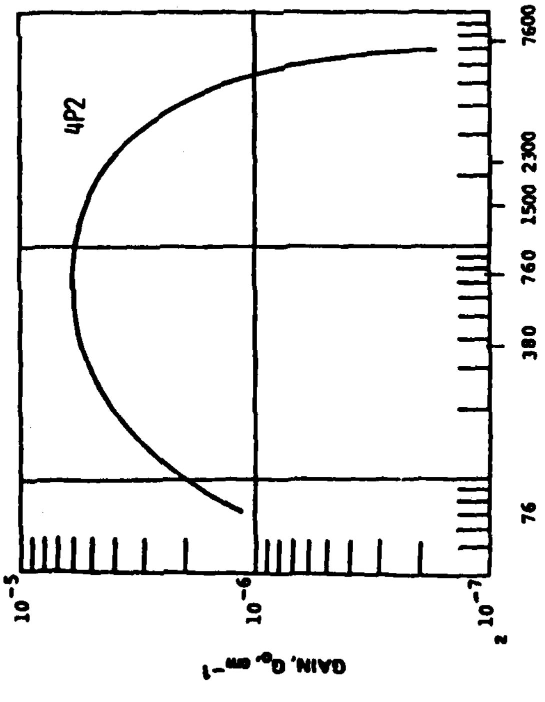
4. $N = g/s = 2.0 \times 10^{-5}\ cm^{-1}/(8.9 \times 10^{22}cm^{2}) = 2.2x10^{16}/cm^{3}$.

The total population at 4. torr and 77°K is $5x10^{16}/cm^{3}$ and the fraction of the gas required to be in the upper state to meet the gain threshold is .04. We assumed, as a working hypothesis based on (4), that we could obtain sufficient population to overcome mirror losses plus about the same amount of gain to build up the laser flux. Under this gain the time to produce 10^{16} photons from one initiating photon is:

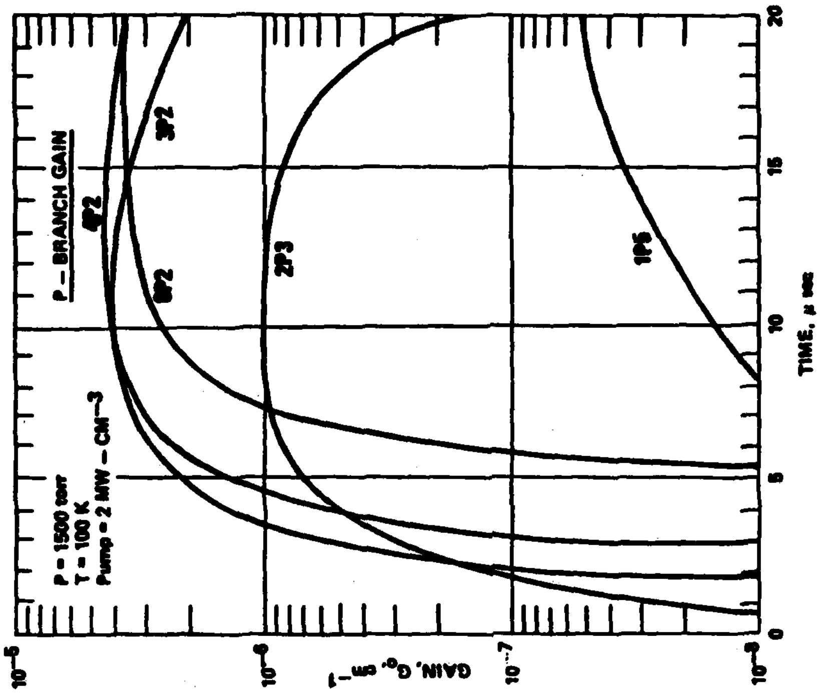
5. $t = \ln(N)/(cxa) = \ln(10^{16}/(2x10^{-5}x3x10^{10}) = 60.$ microsec., where c is the velocity of light and a is the small signal gain.

The low gain requires that a long path, minimum loss laser be constructed to provide the best possible chance for laser action in HD. The laser described later in the text resulted from this requirement. It includes both a long path length, for maximum gain per "round trip," and a hemispherical cavity for minimum diffraction losses.

A detailed kinetics computer program was developed concurrently with the laser experiment. This program explicitly addressed the calculation of gain in the laser as a function of time for various electrical pumping conditions using a finite-rate kinetics analysis (5). The results indicate that small gains are obtainable. However, the magnitude and temporal duration of the gain are not sufficient to meet the gain-time requirements shown in equations, 3 and 5, above. Figure 1. shows the computed maximized gain occurring at 1500. torr (not 4. torr as in our experiment) and figure 2. shows that sufficient gain still falls short of the minimum time to produce a sensible output. With



Fig(1). Pressure Dependence of Gain for 4P2 Transition in HD



Fig(2). Small Signal Gain vs Time for HD these computational results in hand and having not gotten laser output experimentally on HD we immediately turned our attention to making a transfer-storage laser, type (i).

The H_2/HCl and D_2/HCl Experiments: The laser was then converted, by change of mirrors, to run as a type (i) H_2/HCl or D_2/HCl laser. It was in this configuration that the presently reported lasers were first seen. The work on HD showed the sensitivity of the proposed experiments to gain so that working back through the time and threshold gain problem we found the conditions under which laser action could be detected in a few microseconds. The minimum gain for 1 microsec. generation of 10^{16} photons is:

6. $a = \ln(N)/cxt = 36.8/(3\times10^{10}\times1\times10^{-6}) = 1.2\times10^{-3}\text{cm}^{-1}$

Previous work (6) showed that this gain might be achievable. Our experiments produced lasers in the midinfrared, first in H_2/HCl and later in D_2/HCl, but they were both found to be solely associated with H_2 and D_2. We stopped work on H_2/HCl and D_2/HCl to investigate these lasers. As will be explained later, they are newly discovered electronic transition lasers and not storage lasers.

Experiment Design: The experiment (7) is sketched in figure 3. The discharge tube consisted of a 2.5 cm diam Pyrex conical glass pipe. This included two, 3.0 m sections plus end Ts', discharge electrodes, mirror assemblies, and a center 5. cm x 30. cm cross section with an appended cold trap. The cavity output end used a flat sapphire window-mirror, dielectric coated to 95% reflectivity at 3.7 microns. A concave Pyrex mirror, 20 m radius of curvature, dielectric coated to 99+%, formed the other end. Both mirrors were partially transmitting in the visible, and a He-Ne laser was used for alignment of the cavity. Output was detected on a gold-doped germanium, liquid nitrogen-cooled detector. Detector voltage, discharge current, and discharge voltage were recorded on a storage oscilloscope.

The discharge was produced by a Marx bank fired through a resistor and across a gap into the laser tube. The bank consisted of four stages, each a 0.032 microfarads, 125kV, fast discharge capacitor. Normal charge voltage was in the range between 60 and 75 kV with corresponding erected voltages of 240 to 300 kV. Voltage was measured using a 10:1 RC voltage divider with a high voltage oscilloscope probe to provide an overall attenuation of 10000:1. Discharge current was measured by an inductive current transformer in the ground return circuit.

Typical operating conditions were: gas pressure, 2-8 Torr; current, 100-1200 amperes; run voltage (as opposed to initiation voltage) 10-50 kV; discharge duration, 0.5-8 microsec. Current, run voltage, and discharge times were primarily dependent on the discharge series resistors, which were wire wound, ceramic coated and ranged in value from 25-1000 ohms.

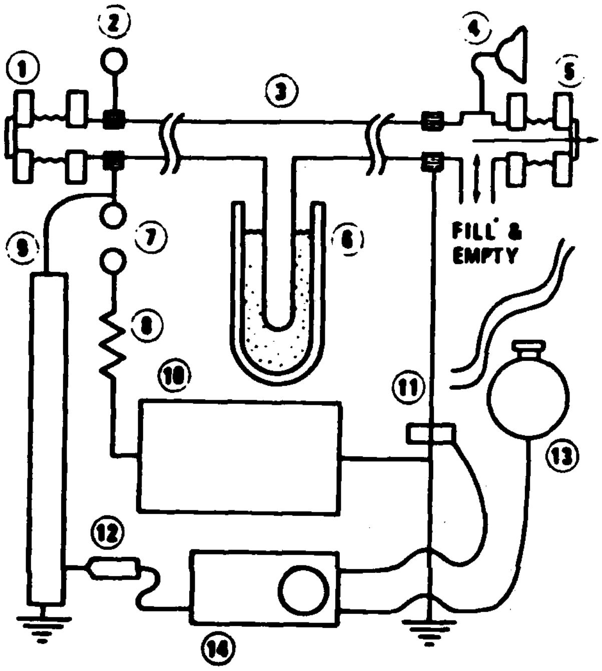


Fig (3). Experimental schematic for the laser: (1) concave mirror on adjustable sealed mount; (2) corona ring for high voltage end; (3) two 3. m sections of 2.5 cm diam Pyrex conical glass pipe; (4) pressure gauge, 0-20Torr; (5) output mirror on adjustable sealed mount; (6) liquid nitrogen cold trap; (7) discharge peaking gap; (8) current limiting resistor; (9) 10:1 voltage divider; (10) four-stage Marx bank; (11) current loop; (12) high voltage probe; (13) Au-Ge detector; and (14) wideband storage oscilloscope.

In the initial phase of the H_2 research, most experiments were done at room temperature; however, some were done with the tube in contact with a liquid nitrogen bath. This gave an estimated gas temperature of 100°K, inferred from pressure change in the tube during cooling. Some changes in laser action were noted during cold tests; however, these may have been caused by alignment problems and no conclusions were drawn from this experiment. The appended cold trap was also tested and after an hour of trapping at liquid nitrogen temperature no difference in laser performance was noted.

A quarter meter grating monochromator with a 150 line/mm grating and an 800 micron slit system was used for coarse location of the lines. The wavelength was measured using a half meter Ebert monochrometer with a 148-lines/mm grating blazed at 5 microns. Slit widths between 400 and 100 microns were used, the former to scan for line center and the latter to obtain the wavelength. Calibration was by comparing dial settings to the 5th, 6th, and 7th order of the He-Ne laser. The wavelength measurements were not sufficiently accurate to include the fine correction for dispersion.

The H_2 Line at 3.767 Micron: - Parametric variation experiments were being performed on a mixture 95% hydrogen/5% hydrogen chloride when the laser was first discovered. Two distinct laser output patterns were noted. During the high current discharge a very short (0.5 microsec.) relatively high intensity (7 volt detector output) pulse close to the leading edge of the current pulse was observed. As the peak current was lowered, i.e., series resistance increased, the pulse occurred later in the current cycle, was lower in peak intensity, and stretched out to several microseconds. Eventually, when the 500 and 1000 ohms resistors were used, and the current was down to 150 amps or less, the pulse took on a distinctively different shape. It appeared after the current peak, followed the current shape for 3-4 microsec. and then terminated. Figure (4) shows oscilloscope traces for a short pulse; Figure (5) shows similar traces for a long pulse.

The question of the source of the laser transition was addressed by trying various mixes of gases, by the cold tube and cold trap experiments, and by removing spurious sources of unknown gases insofar as was possile. Mixtures of hydrogen-nitrogen, hydrogen-helium, and hydrogen-hydrogen chloride had varying effects on the output, but laser action continued until dilution effects dominated. However, a deuterium-hydrogen mixture appeared to suppress the laser action. No significant difference in output was noted between commercial grade and spectroscopic grade hydrogen. Laser action was terminated by cavity detuning; a tuned to detuned intensity ratio was at least 4 orders of magnitude. Particular attention was given to the possibility that the transition might be in hydrogen chloride, since the laser was first seen with a hydrogen-hydrogen chloride mixture in the tube, and only laser transitions known to us near this wavelength are hydrogen chloride and deuterium fluoride transitions. In hydrogen chloride the V(2-1)P(6) vibration-rotation transition lies at 3.77 microns. Spectrometric measurements, however, showed that the wavelength was definitely different from this hydrogen chloride line, and all the experiments mentioned here indicated that only hydrogen was involved in the laser.

Other Lines in H_2 and D_2: The D_2 lines were discovered under conditions similar to those of the H_2 line. In the case of D_2 a mixture of 95% deuterium/5% hydrogen chloride was electrically excited in the laser cavity. No laser action was observed at room temperature; however, because the kinetics would be more favorable at low temperature we cooled the laser with liquid nitrogen and repeated the experiments. Laser action was noted under these circumstances at 4.60 microns. Again, tests were made to determine the source and all but deuterium was eliminated by the same procedure used on hydogen.

Replacement of the mirrors was required when the coating came off the concave cavity mirror. The new mirrors were coated for the 3.7-4.7 micron band with maximum reflectivity on the concave mirror (99+%) and from 85% to 95% on the flat mirror (output end).

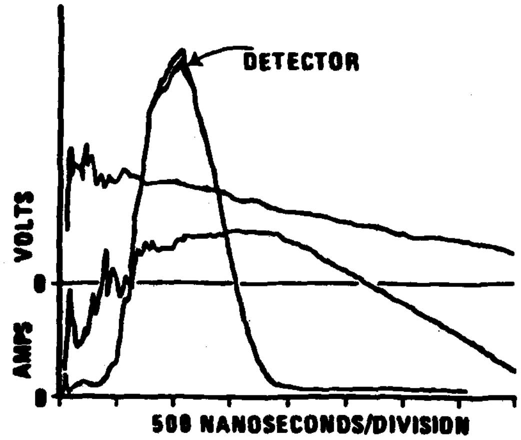


Fig (4). Short pulse, 50 ohm series resistance

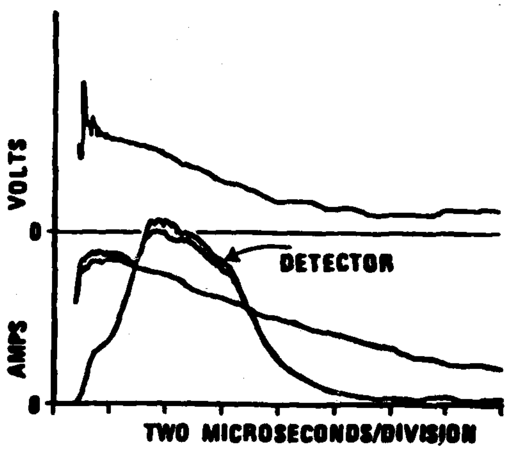


Fig (5). Long pulse, 500 ohm series resistance

All experiments were repeated with the new set of mirrors. Altogether three lines were found in hydrogen and three in deuterium. The hydrogen lines, 3.70 microns, 3.77 microns, and 3.84 microns were obtained at room temperature. The deuterium lines were found only when the laser was cooled with liquid nitrogen. The lines are, in order of decreasing intensity, 4.60 microns, 4.52 microns, and 4.70 microns. No laser lines were found in deuterium hydride at either temperature.

Spectroscopic Assignment of Lines: Following the work of (8), (9) and using the spectroscopic constants and notation of (10) we matched the 3.70 micron line to the singlet transition E-B, V(0-5) P(2) and the 3.77 micron line to the singlet transition K-C, V(1-6) P(4). A kinetically plausible set for all three lines was not found in this series.

Examination of the triplet structure produced entirely different results in which the experimental and calculated wavelengths for hydrogen are closely matched and fall into a kinetically linked pattern in the $a^3\Sigma_g^+ - c^3\Pi_u$ band as follows: 3.70 microns, V(3-2)P(2); 3.77 microns, V(2-1)P(2); 3.84 microns, V(1-0)P(2). The set of calculated lines in this band which fall close to the ones observed in hydrogen are:

TABLE I. Calculated wavelengths in the $a^3\Sigma_g^+$-$c^3\Pi u$ band of H_2

	microns		microns		microns
V(1-0)P(1)	3.76	V(2-1)P(1)	3.69	V(3-2)P(1)	3.63
P(2)	3.84	P(2)	3.77	P(2)	3.71
P(3)	3.93	P(3)	3.85	P(3)	3.78
				P(4)	3.85

The temperature dependence of the rotational state population is given by:

7. $J\ max = .59\ (T/B)^{1/2} - .5$

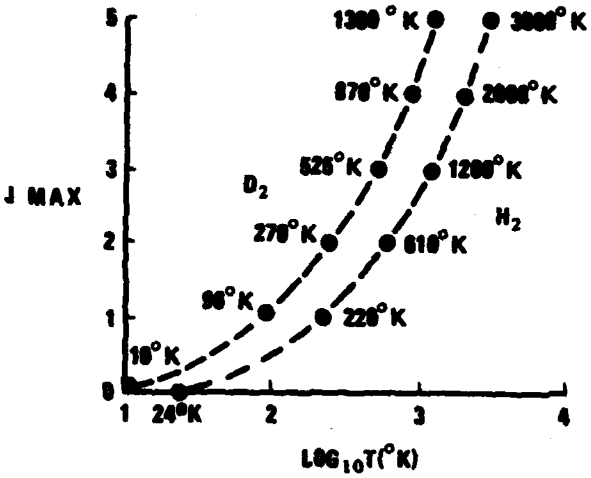
Figure 6. shows the graph of this function for hydrogen in the $a^3\Sigma_g^+$ state which $B = 34.cm^{-1}$ and for deuterium in the $c^3\Pi u$ state in which $B = 15.cm^{-1}$. Reference (10) indicates that the upper laser level in D_2 may be $c^3\Pi u$ state instead of the $a^3\Sigma_g^+$ state. P(2) appears to be the preferred rotational transition for both gases, which matches the experimental observation that hydrogen produces lasers at 300°K and 100°K. Experimentally, we find that the V(2-1) transition in H_2, the strongest line, occurs first and is followed in about 0.2 microsec. by the other two lines. This is not unusual in cascading lasers; it is an indication that the laser action on the main line kinetically produces conditions for the other two lines to appear. Good computational modeling should show this effect, too. This delay is shown in the oscillograph traces of Figure 7.

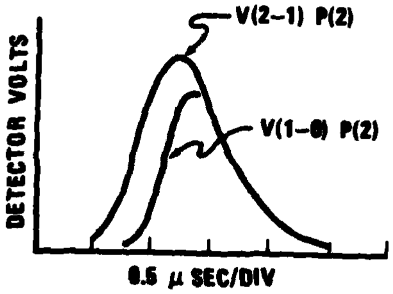
An identical assignment cannot be made for deuterium: first, because its $c^3\Pi u$ state, lies above the $a^3\Sigma_g^+$ state and second, because the vibrational and rotational coefficients (except B) are not known. In principle the laser lines discovered in deuterium could be used to produce spectroscopic coefficients. This will be tried at a later date, particularly if additional lines can be found and more accurate wavelenghts determined.

Cavity Properties: The laser cavity composed of a plane output mirror and a concave back mirror, figure 8, may be considered to be half of a symmetric resonator. It has an effective mirror diameter of 2.5 cm and a full cavity length of 1260 cm (twice the mirror spacing). A typical wavelength is $4.x10^{-4}$ cm so the Fresnel number is approximately 3. This configuration thus has negligible diffraction loss. The resonator g parameter is:

8. $g = g1 = g2 = 1\text{-}L/R = 1\text{-}1260/2000 = .37$

and the stability factor, g^2 is 0.137, signifying a stable oscillator.





Fig(6). J max for H_2, $a^3\Sigma_g^+$ and for D_2, $c^3\Pi_u$

Fig(7). V(1-0)P(2) line delay with respect to V(2-1)P(2)

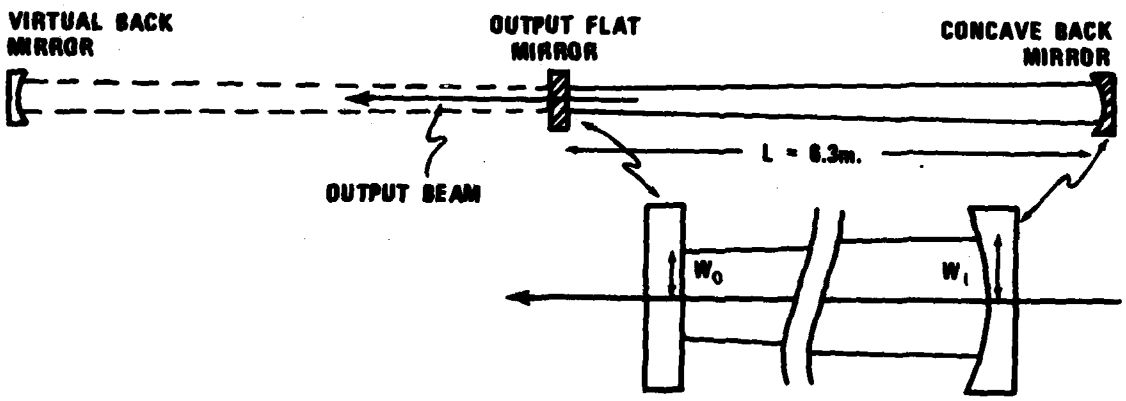

VIRTUAL BACK MIRROR
OUTPUT FLAT MIRROR
CONCAVE BACK MIRROR
OUTPUT BEAM
L = 6.3m.
W₀
W₁

Fig. 8. Equivalent Cavity

The beam waist, radius w_0, is at the flat output mirror and is,

9. $w_0 = (2L/\Pi)^{\frac{1}{2}} (g^2(1-g^2))^{\frac{1}{4}}/(2g-2g^2)^{\frac{1}{2}} = .35$ cm

The back mirror spot size, w_1, is

10. $w_1 = (2L/\pi)^{\frac{1}{2}} (1-g^2)^{-\frac{1}{4}} = .38$ cm

The average mode radius is .37 cm as compared to the discharge radius of 1.cm; thus the mode fills above .14 of the discharge volume.

Excitation Discharge Properties: The general circuit block diagram is shown in Figure 3. Properties of the discharge were calculated from the measured current and voltage and from previously calculated (4) transport properties of hydrogen and deuterium discharges. Energy distribution to vibration, rotation, dissociation and translation were determined as functions of the input parameters: pressure, temperature, voltage and current. Discharge properties are determined for hydrogen as follows: given current and voltage as functions of time as shown in Figure 9., we calculate the E/N values as functions of time. Drift velocities, Wd, for the electrons are then taken from the curves of Wd vs E/N (4). The free electron concentration and ion concentrations are calculated from these data by:

11. Ne (=Ni) = I/(qxWdxA)

where I is the measured current, q is the charge on the electron, and A is the cross section of the discharge. Likewise power input and the distribution of power to the various modes were tabulated as functions of time from the graphs of (6) which show energy distribution as obtained from a Boltzmann calculation.

The rate of electron-ion pair formation and the concentration of these are thus determined from measured and computed discharge properties. The properties associated with the discharge shown in Figure 9. are given in Table II below:

Table II. Properties of a Typical Discharge in Hydrogen

TIME	POTENTIAL	CURRENT	POWER	ENERGY	EL.DENS.	DIST. OF POWER		
(usec)	(kV)	(amps.)	(kw/cm^3)	(joules/cm^3)	cm^{-3}	pd	pv	pT
0.5	110.	1200.	64.	.017	1.8×10^{14}	.70*	.18*	.07*
1.0	80.	1240.	48	.045	2.7×10^{14}	.51	.29	.13
1.5	52.	910.	23	.062	2.4×10^{14}	.32	.44	.20
2.0	35.	580.	9.8	.069	2.4×10^{14}	.05	.64	.27
2.5	25.	300.	3.6	.075	1.8×10^{14}	.0	.66	.31

*estimated by extrapolation

Distribution of power terms represent the power at any time going into the various uses as follows: Pd into dissociation, Pv into vibrational excitation, and PT into thermal plus rotational excitation

On-set of laser action typically occurred, for the discharge with a 50 ohms series load, at 1. microsec. The gas properties at that time are noted in the Table II. Specific energy loading in the hydrogen at 1.microsec is 8500. joules/liter atmos.

This may be compared to the best loading for e-beam CO_2 lasers at one atmosphere which is about 900 joules/liter atmos. The eventual loading on the hydrogen in this example was 14,000. joules/liter atmos. at 2.5 microsec.

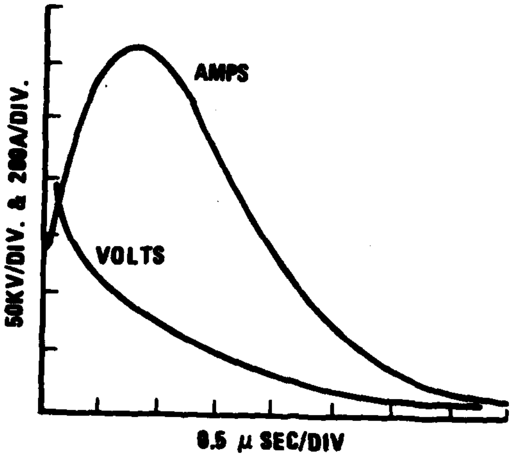


Fig. 9. Oscillograph trace of current and voltage for the experimental laser (Fig. 3.) when the series load was 50 ohms.

CONCLUSION: The research program on hydrogen as a storage laser for high energy lasers applications has shown that HD certainly cannot be made into a storage laser under the operating conditions we had and perhaps cannot be so made under any circumstance. No conclusion was reached concerning H_2/HCl or D_2/HCl since serendipity produced a new set of lasers in the course of experiments on these gas mixtures. The new lasers, identified at least in hydrogen as being in the triplet states, may be the first step in successful demonstration of the hydrogen excimer (11).

REFERENCES:

1. N. G. Basov, A. N. Oraevskii, and A. F. Suchkov, JETP Lett. 16, 211. (1972)

2. D. W. Howgate, T. G. Roberts, and E. T. Gerry, "New Laser Concepts-Executive Summary Report," TR H-78-2 US Army Missile Command. (Nov 1978)

3. A. R. W. McKellar, Can. J. Phys. 51 389, (1973); Astrophys. J. 185, L 53 (1973); Can. J. Phys. 52, 1144. (1974)

4. A. G. Engelhardt and A. V. Phelps, Phys. Rev. 131, 2115. (1963)

5. T. A. Barr, Jr., D. W. Howgate, W. B. McKnight, and R. L. Taylor, "Diatomic Molecules as Storage Media for High Energy Lasers," International Conference, Lasers 81, New Orleans. (Dec. 1981)

6. R. Taylor, "Analytical Modeling of Electrically Excited D_2/HCl and HCl Laser Experiments," PSI TR-58 Final Rpt. on Contr. # N00014-75-C-0035. (July 76)

7. W. B. McKnight and T. A. Barr, Jr., Appl. Opt. 21, No 3. 357-358. (1982)

8. Handbook of Lasers, ed. by R. J. Pressley, 319-322, The Chemical Rubber Company, Cleveland, Ohio. (1971)

9. P. A. Bazhulin, I. N. Kruyazev, and G. G. Petrash, Sov Phys JETP, 20, 1068-1069 (1965) and Sov Phys JETP, 21 11-16. (1966)

10. K. P. Huber and G. Herzberg, Molecular Spectra and Molecular Structure, IV, Constants of Diatomic Molecules, 240-251, 262-267, Van Nostrand Reinhold Company, New York. (1979)

11. C. V. Heer, Phys. Lett. 31A, No. 4, 160-161 (1970)

Gravity Vector Determination from Inertial and Auxiliary Data and Potential Utilization of Generated Vector Component Information

H. BAUSSUS VON LUETZOW, DR.
U.S. Army Engineer Topographic Laboratories
Fort Belvoir, Virginia 22060

1. INTRODUCTION. The present Rapid Geodetic Survey System (RGSS), developed by Litton Systems, Inc. for the U.S. Army Engineer Topographic Laboratories (ETL), has the capability of determining horizontal and vertical positions, deflections of the vertical, and gravity anomalies with average rms errors of 1 m, 0.3 m, 2 arcsec, and 2 mgal, respectively, for 50 km runs under utilization of post-mission adjustments. It operates as a quasi local-level system. It thus does not require altitude damping, permits Kalman stochastic error control without great complexity under consideration of observed velocity errors at vehicle stops, and employs effective post-mission adjustments with the aid of terminal position and gravity vector information only. Present critical hardware consists of an A-1000 vertical accelerometer, two A-200 horizontal accelerometers, and two G-300 gyroscopes. Although data utilization from repeated runs has resulted in average deflection rms errors of about 1 arcsec, the requirement of a maximum rms error of 0.5 arcsec necessitates the incorporation of higher performance gyroscopes and accelerometers. This would simultaneously achieve conventional surveying accuracy (10^{-5}). Following the approximate elimination of the effects of constant gyro biases on positions under utilization of accurate initial and terminal coordinates, the remaining dominant error source is gyro correlated random noise. In this respect, the pertinent parameters for the G-300 instruments are a standard deviation of 0.002° hr^{-1} and a correlation time of 3 hours. Another error source to be considered is the accelerometer scale factor, assumed to be constant for a test run. The standard scale factor error for the A-200 accelerometer is 0.01%. Correlated accelerometer noise is characterized by a standard deviation of 10 mgal and a correlation time of 40 minutes. Both correlated gyro random drift and correlated accelerometer measurement errors affect the accuracy of deflections of the vertical estimated from inertial data and initial and terminal deflection components, possibly augmented by corresponding geodetic azimuths. In order to achieve

deflection accuracies smaller than or equal to 0.5 arcsec rms, it is, therefore, necessary to install higher performance gyros and accelerometers. A significant reduction of the two autocorrelation parameters is also a necessary prerequisite for the optimal statistical estimation of deflections of the vertical. The Litton G-1200 gyro is expected to be compatible with the stringent deflection accuracy requirement. Its random drift error is about $0.001^\circ\ hr^{-1}$ and its autocorrelation parameter is short. For details, reference is made to Litton [1] and Huddle [2]. The A-1000 accelerometer addressed by Litton [3,4] has a standard scale factor error of approximately 0.005%. Correlated accelerometer noise is 2 mgal rms, equivalent to about 0.35 arcsec.

2. DETERMINATION OF VERTICAL GRAVITY COMPONENT. In a local Cartesian coordinate system tangent to a normal gravity surface, with x as west-east coordinate, y as south-north coordinate and z as vertical coordinate, positive upward, it is

$$\vec{g} = (g\eta,\ -g\xi,\ -g) \qquad (1)$$

In eq. (1), $\vec{g}$ is the gravity vector at the origin, $-g$ is its vertical component, and $g\eta$ and $-g\xi$ are its horizontal components, η and ξ being the prime and meridian deflections, respectively. For computational and estimation purposes it is advantageous to use the gravity anomaly $\Delta g = g - \gamma$ with γ as normal gravity, to be found by consideration of the elevation above the normal reference surface.

The RGSS error differential equation for the vertical channel, including first and second order time derivatives, is

$$\ddot{z} = 2\omega_s^2 z - (\rho_N + 2\,\Omega_N)\,\dot{x} - \rho_E\,\dot{y} + \Delta g + a_z + A_x\delta_{zx} - A_y\,\delta_{zy}\ [1] \qquad (2)$$

where the symbol δ of position errors δx, δy, δz is omitted for simplicity. In eq. (2), $\omega_s^2 = g\,R^{-1}$ with R as the earth's mean radius is the squared Schuler frequency, $\rho_N = V_x R^{-1}$ with V_x as the system's east velocity is the north angular rate, Ω_N is the north earth rate, $\rho_E = -V_y R^{-1}$ with V_y as the system's north velocity is the east angular rate, a_z is the vertical accelerometer error, A_x and A_y are east and north accelerometer outputs, and δ_{xz} and δ_{yz} are z-axis misalignment errors in the x,z and y,z planes. At the system's stop, the term involving $\dot{x}$ may be estimated since x is observed, and the last two terms are omitted. At terrestrial vehicle speeds, the term involving $\dot{y}$ becomes insignificant.

---

[1] See Litton [5], p. 3-132.

Because of initial calibration $\ddot{z}(t_0) = \ddot{z}_0 = 0$ and $z_0 = \dot{x}_0 = \dot{y}_0 = A_{xo} = A_{yo} = 0$, the term $\Delta g_0 + a_{z0}$ has to be subtracted from $\ddot{z}(t_\nu) = \ddot{z}_\nu$. At a survey vehicle stop, therefore, under omission of $\rho_N \dot{x}$ and $\rho_E \dot{y}$ in eq. (2),

$$\Delta g_\nu = \Delta g_0 + \ddot{z}_\nu + 2\Omega_N \dot{x}_\nu - 2\omega_s^2 z_\nu - (a_{z_\nu} - a_{z_0}), \tag{3}$$

where the error terms $2\omega_s^2 z_\nu$ and $a_{z_\nu} - a_{z_0}$ are weakly correlated. If z and a_z are considered as random variables with zero means and estimated variances and covariances, respectively, the gravity anomaly error variance is

$$\text{var}\ \delta g_\nu = 4\omega_s^4\ \text{var}\ z_\nu + \text{var}\ (a_{z_\nu} - a_{z_0}) \tag{4}$$

Since z_ν has an average rms error of 0.3 m due to Kalman filter estimation by means of observed z_ν at vehicle stops and a post-mission adjustment under utilization of a terminal z_e, the rms error induced by z_ν is only 0.1 mgal. High accuracy determinations of gravity anomalies or of the vertical component of gravity by means of a local-level inertial system thus require a high performance vertical accelerometer, high accuracy height determinations and, possibly, utilization of $\dot{x}$-observations at vehicle stops.

The availability of Δg_e and z_e at the end of the survey permits the determination of $a_{z_e} - a_{z_0} = f_e$ in eq. (3). Empirical corrections for $a_{z_\nu} - a_{z_0}$ in the form

$$C_\nu = c_\nu (a_{z_e} - a_{z_0}) = c_\nu f_e \tag{5}$$

prevent error growths to about 2 var a_z. Under restrictive assumptions, with the omission of the subscript z,

$$c_\nu = \text{cov}\ (a_\nu - a_0, a_e - a_0) \cdot [\text{var}\ (a_e - a_0)]^{-1} \tag{6}$$

3. SINGLE HORIZONTAL CHANNEL DETERMINATION OF DEFLECTIONS OF THE VERTICAL. **While the determination of gravity anomalies may be less rapidly accomplished with high accuracy by means of gravimeter and height measurements, astrogeodetic and gravimetric determinations of**

deflections of the vertical are time-consuming. According to Moritz [6], astrogeodetic and gravimetric rms errors are approximately 0.2 arcsec and 1.5 arcsec, respectively. Rapid inertial-astronomic determinations of η and ξ with rms errors not exceeding 0.35 arcsec would thus be highly economical and satisfactory for many purposes. Both the single and coupled horizontal channel determinations of η and ξ require the utilization of the simplified error differential equations[2]:

$$\ddot{x} = S_N\phi_z - g\phi_N + g\eta + a_E \tag{7}$$

$$\ddot{y} = -S_E\phi_z + g\phi_E - g\xi + a_N \tag{8}$$

$$\dot{\phi}_z = R^{-1}\text{tn}\phi.\dot{x} + R^{-1}(\Omega_N + \rho_N \sec^2\phi)y \qquad -\omega_E\phi_N \quad +\omega_N\phi_E + \alpha \tag{9}$$

$$\dot{\phi}_N = R^{-1}\dot{x} \qquad +\omega_E\phi_z \qquad -\omega_z\phi_E + \beta \tag{10}$$

$$\dot{\phi}_E = -R^{-1}\dot{y} \qquad -\omega_N\phi_z \quad +\omega_z\phi_N \qquad +\gamma \tag{11}$$

Symbols used in the foregoing equations including total time derivatives of first and second order are, except for those defined in section 2,

ϕ_z	azimuth platform attitude error
ϕ_N	platform tilt error about north axis
ϕ_E	platform tilt error about east axis
S_E	east acceleration of survey vehicle
S_N	north acceleration of survey vehicle
a_E	correlated east accelerometer error
a_N	correlated north accelerometer error
ϕ	geographic latitude
ω_z	$= \Omega_z + R^{-1}\text{tn}\,\phi\, V_x$ vertical spatial rate
ω_N	$= \Omega_N + \rho_N$ north spatial rate
ω_E	$= -R^{-1}V_y$ east spatial rate
α	azimuth axis angular drift rate error
β	north axis angular drift rate error
γ	east axis angular drift rate error

For land vehicles, eqs. (9) - (11) may be simplified by omission of ρ_N,

---

[2] A more complex system may be required. See Litton [5], pp. 3-252 - 3-254.

ω_E, and $\rho_z = R^{-1} tn\phi . V_x$, and by use of constant accelerations S_N and S_E which should be approximately achieved. Then, $\omega_z = \Omega \sin\phi$ and $\omega_N = \Omega \cos\phi$ where Ω denotes the earth's inertial angular velocity. The initial conditions at $t_o = 0$ are, under consideration of plumbline leveling, $\omega_z(0) = 0$, $\phi_N(0) = \eta_o$, $\phi_E(0) = \xi_o$, $x(o) = y(o) = x(o) = y(o) = 0$.

In the general RGSS mode of operation, position and gravity vector information is generated. The Kalman filter mechanization is simplified by setting $\omega_E = \omega_N = \omega_z = 0$ in eqs. (9) - (11) which results in the uncoupling of the horizontal channel differential equations. This is not critical in the context of present instrumentation for travel periods of 2 hours. As an example, eqs. (8) and (10) may then be formulated at the end of the first travel period, whence $t = t_1$,

$$\ddot{y}_1 = g\phi_1 - g\xi_1 + a_{N_1} - a_{N_o} \tag{12}$$

$$\phi_{E_1} = -R^{-1} y + \int_o^{t_1} \gamma dt + \xi_o \tag{13}$$

In eq. (12), accelerometer calibration at $t_o = 0$ has been taken into account. The quantity $-R^{-1}y$ in eq. (13) is well estimated by the observed velocity error y. After implementation of the corresponding tilt correction, under neglect of insignificant estimation errors and consideration of the bias error $\bar{\phi}_E(t_i)$,

$$\ddot{y}(t_i) = \ddot{y}_i = g\bar{\phi}_{E_i} - g\,(\xi_i - \xi_o) + g \int_o^{t_i} (\gamma - \bar{\gamma})\, dt + a_{N_i} - a_{N_o} \tag{14}$$

$$\ddot{y}(t_n) = \ddot{y}_n = g\bar{\phi}_{E_n} - g\,(\xi_n - \xi_o) + g \int_o^{t_n} (\gamma - \bar{\gamma})\, dt + a_{N_n} - a_{N_o} \tag{15}$$

The propagation of constant gyro bias errors may be approximated by a solution of eqs. (9) - (11) with $y = x = y = \omega_E = 0$ and by treating ω_N and ω_Z as constants in accordance with Huddle [8]. The result is

$$\bar{\phi}_E = t\,(\bar{\gamma} \qquad + \bar{\beta} t\Omega_z/2 \; - \bar{\alpha} t\Omega_N/2) \tag{16}$$

$$\bar{\phi}_N = t\,(- \bar{\gamma} t\Omega_Z/2 \quad + \; \bar{\beta} \qquad) \tag{17}$$

$$\bar{\phi}_z = t\,(\bar{\gamma} t\Omega_N/2 \qquad + \bar{\alpha} \quad) \tag{18}$$

Under consideration of terminal deflection and azimuth errors at $t = t_n$

it is possible to estimate $\bar{\alpha}$, $\bar{\beta}$, $\bar{\gamma}$ from

$$\begin{bmatrix} \bar{\phi}_E \\ \bar{\phi}_N \\ \bar{\phi}_z \end{bmatrix}_{t_n} = \begin{bmatrix} \xi_n \\ \eta_n \\ A_n \end{bmatrix}_{\text{estimated}} - \begin{bmatrix} \xi_n \\ \eta_n \\ A_n \end{bmatrix}_{\text{observed}} \tag{19}$$

Equations (15) and (16) show that the computation of $\bar{\phi}_{E_i}$ in eq. (14) is to the first order associated with an error

$$\delta\bar{\phi}_{E_i} = - \frac{t_i}{t_n} \left[g \int_o^{t_n} (\gamma - \bar{\gamma})\, dt + a_{N_n} - a_{N_o} \right] \tag{20}$$

The final solution for the meridan deflection of the vertical from eqs. (14) and (20) is

$$\xi_i = \left(\xi_o + \bar{\phi}_{E_i} - \frac{\ddot{y}_i}{g} \right) + \frac{a_{N_i} - a_{N_o}}{g} - \frac{t_i}{t_n} \frac{a_{N_n} - a_{N_o}}{g} + \int_o^{t_i} (\gamma - \bar{\gamma}) dt - \frac{t_i}{t_n} \int_o^{t_n} (\gamma - \bar{\gamma}) dt \tag{21}$$

Equation (21) may be supplemented by

$$\delta\xi_i = \delta\xi_o + \frac{t_i}{t_n} (\delta\xi_e - \delta\xi_o) - \left(\Phi_{S_i} - \frac{t_i}{t_n} \Phi_{S_n} \right) \tag{22}$$

to account for astrogeodetic deflection errors and accelerometer bias errors. For a straight traverse, the last term in eq. (22) tends to cancel out.

The rms deflection error $\sigma_\xi(t_i)$ can be computed by covariance analysis involving the terms without parentheses in eq. (21). Under inclusion of the first two terms of eq. (22),

$$\text{var}\, \xi_i = \text{var}\, a_i + \text{var}\, \gamma_i + \left(1 - \frac{t_i}{t_n}\right)^2 \text{var}\, \xi_o + \left(\frac{t_i}{t_n}\right)^2 \text{var}\, \xi_n , \tag{23}$$

where var a_i is the accelerometer-induced variance and var γ_i designates the gyro-induced variance. Under consideration of present RGSS parameters, identified in section 1, the rms deflection error σ_{ξ_i} is approximately presented in Figure 1 for a vehicle speed of 30 km/hr

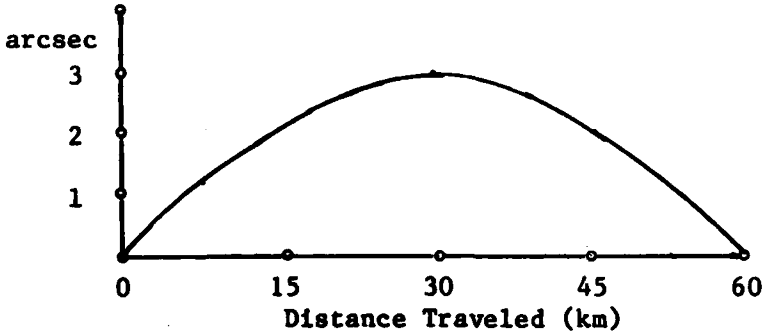


Figure 1
Approximate RMS Deflection Error as a Function of Travel Distance

The variation of normal gravity g with altitude is not critical in terrestrial applications. The scale factor-generated error does not vanish with respect to L-shaped traverses. It is, however, relatively small for travel times not exceeding 2 hours. Heading sensitivity induced by significant azimuth changes appears to be somewhat critical and may require empirical corrections. Schwarz [7] discussed the impact of heading sensitivity on positioning and concluded that error reductions may be achieved by a modified smoothing procedure. The problem is more intricate in the case of deflection determinations. The use of static accelerometer measurements in conjunction with initial and terminal deflections together with a simplified Kalman filter was first discussed by Huddle [8]. An advanced method for single channel vertical deflection determination was presented by Baussus von Luetzow [9]. It utilizes all observations simultaneously and appears in a modified form in section 4.

4. COUPLED HORIZONTAL CHANNEL OPTIMAL DETERMINATION OF DEFLECTIONS OF THE VERTICAL IN SEMI-FLAT TERRAIN. Higher accuracy requirements necessitating installation of gryoscopes and accelerometers with small error variances and short correlation times as performance characteristics require the integration of the system of differential equations for best possible estimation of η and ξ and for error assessment in the case of single channel solutions. For this purpose, eqs. (7) and (8) are written as

$$\ddot{x} = S_N\phi_z - g\phi_N + g(\eta - \eta_o) + a_E - a_{E_o} \tag{24}$$

$$\ddot{y} = -S_E\phi_z + g\phi_E - g(\xi - \xi_o) + a_N - a_{N_o} \tag{25}$$

In this way, initial accelerometer calibrations have been taken into account and the initial conditions are

$$\phi_E(o) = \phi_N(o) = \phi_z(o) = \dot{x}(o) = \dot{y}(o) = x(o) = y(o) = 0 \tag{26}$$

The accelerations S_N and S_E can probably be neglected. If the system is treated as one with constant coefficients as a good approximation, a closed solution as a function of time is possible. Because of intermittent Kalman filter corrections and the need for numerical weight factors it appears to be advantageous to attempt a numerical solution for x and y under utilization of terminal deflection and azimuth data. For economic reasons it is assumed that the survey vehicle travels approximately at a constant speed when in motion and stops every 3 minutes for 1 minute. The speed should not exceed 10 msec^{-1} in order to restrict the length of travel intervals.

Solutions for ϕ_E, ϕ_N, ϕ_z, $\dot{x}$, $\dot{y}$, x, y are obtained in accordance with the integration schemes

$$F_1 = F_o + \dot{F}_o \Delta t \tag{27}$$

$$F_\nu = 2F_{\nu-1} - F_{\nu-2} + (\Delta t)^2 \dot{F}_{\nu-1} \tag{28}$$

with $\Delta t = 30$ sec and possibly 60 sec. The solution structure at the end of the first stop interval, indicated by the subscript s=1, is

$$F_{\nu 1} = \sum_1^\nu a_\nu v_\nu + \sum_1^\nu b_\nu w_\nu + \sum_1^\nu c_\nu \alpha_\nu + \sum_1^\nu d_\nu \beta_\nu + \sum_1^\nu e_\nu \gamma_\nu \tag{29}$$

In eq. (29), $v = g(\eta - \eta_o) + a_E - a_{E_o}$, $w = -g(\xi - \xi_o) + a_N - a_{N_o}$.

Under consideration of Kalman filter tilt corrections, assumed for simplicity to eliminate the integrated first two terms in eq. (9) and the first term in both eq. (10) and (11),

$$F_{\nu 1}^{(2)} = \sum_1^\nu A_\nu v_\nu + \sum_1^\nu B_\nu w_\nu + \sum_1^\nu C_\nu \alpha_\nu + \sum_1^\nu D_\nu \beta_\nu + \sum_1^\nu E_\nu \gamma_\nu + T_1, \tag{30}$$

where T_1 and subsequent T_s represents an aggregate of tilt-induced random errors. During the stop interval, eqs. (9)-(11) are integrated with $x = y = y = 0$ with a resultant effect on eqs. (7) and (8). The integration is then continued and yields under utilisation of average biases the solution structure

$$F_{\nu s}^{(2)} = \sum_{\nu=1}^{\nu_s} A_\nu v_\nu + \sum_{\nu=1}^{\nu_s} B_\nu w_\nu + \hat{C}_\nu\bar{\alpha} + \hat{D}_\nu\bar{\beta} + \hat{E}_\nu\bar{\gamma} + R_{\nu s} + T_s \qquad (31)$$

where $R_{\nu s}$ designates residual, random-type terms including $\bar{\alpha} - \alpha_\nu$, $\bar{\beta} - \beta_\nu$, and $\bar{\gamma} - \gamma_\nu$.

At the termination of the survey, when η_n, ξ_n, and A_n are available, it is possible to solve for the gyro biases in the form

$$\begin{bmatrix} \bar{\alpha} \\ \bar{\beta} \\ \bar{\gamma} \end{bmatrix} = \begin{bmatrix} F_1 \\ F_2 \\ F_3 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix} + \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \end{bmatrix} \qquad (32)$$

where the F's are computable, the ε's are an aggregate of v and w-errors, and the δ's are aggregates of gyro random and tilt errors. Substitution of $\bar{\alpha}$, $\bar{\beta}$, $\bar{\gamma}$ from eq. (32) in eq. (31) and separation of v and w into deflection and accelerometer errors leads, under restriction to solutions for x and y, to the final results

$$\eta_s + \sum_{\nu=0}^{n} k_\nu \eta_\nu + \sum_{\nu=0}^{n} l_\nu \xi_\nu = \eta_o + g^{-1}\ddot{x}_s^{(2)} + \bar{\phi}_{N_s}^{(2)} + r_{gs} + r_{as} + r_{ts} + r_{ds} \qquad (33)$$

$$\xi_s + \sum_{\nu=0}^{n} m_\nu \eta_\nu + \sum_{\nu=0}^{n} n_\nu \xi_\nu = \xi_o - g^{-1}\ddot{y}_s^{(2)} + \bar{\phi}_{Es}^{(2)} + \rho_{gs} + \rho_{as} + \rho_{ts} + \rho_{ds} \qquad (34)$$

The last 4 terms on the right side of both eqs. (33) and (34) are aggregates of random errors associated with gyros, accelerometers, tilt corrections, and initial and terminal deflection errors. Equations (33) and (34) are reformulated as

$$\eta_s + \sum_{\nu=0}^{n} k_\nu \eta_\nu + \sum_{\nu=0}^{n} l_\nu \xi_\nu = S_{\eta_s} = M_{\eta_s} - N_{\eta_s} \qquad (35)$$

$$\xi_s + \sum_{\nu=0}^{n} m_\nu \hat{\eta}_\nu + \sum_{\nu=0}^{n} n_\nu \xi_\nu = S_{\xi_s} = M_{\xi_s} - N_{\xi_s} \qquad (36)$$

where S, M, N denote signal, measureable message, and non-measurable noise, respectively.

Under utilisation of deflection of the vertical covariance functions, the prime deflection of the vertical can be optimally estimated in the form

$$\eta_e = A_{ie}\left[S_{\eta_i} + N_{\eta_i}\right] + B_{ie}\left[S_{\xi_i} + N_{\xi_i}\right] \tag{37}$$

where $i = s$, A_{ie} and B_{ie} are matrices of regression coefficients, and the terms in brackets are message matrices. With $k = 0, 1, \ldots, n$, A_{ie} and B_{ie} can be determined from the equations

$$\overline{\eta_e\left[S_{\eta_k}\right]} = A_{ie}\left[\overline{S_{\eta_i}\,S_{\eta_k}} + \overline{N_{\eta_i}\,N_{\eta_k}}\right] + B_{ie}\left[\overline{S_{\xi_i}\,S_{\eta_k}} + \overline{N_{\xi_i}\,N_{\eta_k}}\right] \tag{38}$$

$$\overline{\eta_e\left[S_{\xi_k}\right]} = A_{ie}\left[\overline{S_{\xi_i}\,S_{\eta_k}} + \overline{N_{\xi_i}\,N_{\eta_k}}\right] + B_{ie}\left[\overline{S_{\xi_i}\,S_{\xi_k}} + \overline{N_{\xi_i}\,N_{\xi_k}}\right] \tag{39}$$

In eqs. (38) and (39), where the bar symbol stands for covariance, the noise covariances need only be computed once.

Simplified solutions, particularly in the case of approximately straight traverses, are possible. The coupling of eqs. (33) and (34) should, however, always be considered. If averaged message-type data from repeated surveys are employed, the instrument-generated noise covariances in eqs. (38) and (39) are to be reduced.

5. INTERPOLATION OF GRAVITY ANOMALIES AND DEFLECTIONS OF THE VERTICAL IN MOUNTAINOUS TERRAIN.[3] It is well known that an accurate analytical representation of free-air anomalies in pronounced mountainous terrain can only be achieved by a polynomial of high degree by means of Δg-data available in a network of high resolution. As a consequence, satisfactory linear interpolation requires small mesh sizes Δx, Δy. It is possible to write the gravity anomaly as

$$\Delta g = \Delta g_i + C_t + r \tag{40}$$

where Δg_i is the isostatic anomaly valid for the compensated geoid, C_t represents the aggregate of terms computable from the known topography, and r is a random-type error. In a more general form, also applicable to the optimal estimation of vertical deflections, eq. (40) is reformulated as

$$m = s + n + r\,. \tag{41}$$

---

[3] For details see Baussus von Luetzow [10]

In this equation, m is a "message" variable, s is a "signal" variable, n is deterministic or computable "noise," and r is random-type noise.

Under consideration of a linear signal estimation structure, a signal can then be optimally estimated as

$$\hat{s}_e = L(m_i - n_i - r_i) \tag{42}$$

where L denotes a linear operator and the subscripts e and i refer to the estimation point P_e and measurement points P_i, respectively. The optimal measurement at P_e results in

$$\hat{m}_e = \hat{s}_e + n_e + r_e = L(m_i - n_i) + r_e - L\,(r_i) + n_e\ . \tag{43}$$

The estimation error is

$$e(\hat{m}_e) = e(\hat{s}_e) + e\,[r_e - L(r_i)] \tag{44}$$

The corresponding estimation error resulting from the utilization of topographically unmodified measurements m_i is

$$\begin{aligned} e(m_e) &= \hat{s}_e - L(s_i) + r_e - L(r_i) + n_e - L(n_i) \\ &= e(\hat{s}_e) + e[r_e - L(r_i)] + n_e - L(n_i). \end{aligned} \tag{45}$$

Comparison of eq. (45) with eq. (44) shows that the non-optimal interpolation process is associated with a "topographic" estimation error $n_e - L(n_i)$ which becomes in general intolerable in moderate to rough mountainous terrain and thus induces the requirement of a fine mesh data grid.

Baussus von Luetzow [10] obtained the following solution for vertical deflections:

$$\begin{aligned} \begin{Bmatrix} \xi \\ \eta \end{Bmatrix} = {} & \frac{R}{4\pi G} \iint_\sigma \left[\Delta\hat{g} + G_1\,(\Delta\hat{g})\right] \begin{Bmatrix} \cos\alpha \\ \sin\alpha \end{Bmatrix} \frac{dS(\psi)}{d\psi}\, d\sigma - \frac{\Delta\hat{g} + G_1(\Delta\hat{g})}{G} \begin{Bmatrix} \operatorname{tn}\beta_1 \\ \operatorname{tn}\beta_2 \end{Bmatrix} \\ & + \frac{R}{4\pi G} \iint_\sigma \left[\delta g_1 + \delta g_2 + G_1(\delta\hat{g}_3)\right] \begin{Bmatrix} \cos\alpha \\ \sin\alpha \end{Bmatrix} \frac{dS(\psi)}{d(\psi)}\, d\sigma + \begin{Bmatrix} \delta\xi \\ \delta\eta \end{Bmatrix} \\ & + \frac{C - \delta g_2 - G_1(\delta\hat{g}_3)}{G} \begin{Bmatrix} \operatorname{tn}\beta_1 \\ \operatorname{tn}\beta_2 \end{Bmatrix} \end{aligned} \tag{46}$$

In this equation, G denotes global mean gravity, $\Delta\hat{g}$ is a signal variable profitably synonymous with Δg_i in eq. (40), α is the azimuth angle counted clockwise from north, ψ is the angle between the radius vectors $\vec{r}_A$ and $\vec{r}_p$ originating at the earth's spherical center in accordance with figure 2, S (ψ) is Stokes' function, β_1 and β_2 are northern and eastern terrain inclinations, respectively, G_1 is a specified integral function, δg_1, δg_2, δg_3 are quantities computable from the earth's topography, C is the terrain correction, and σ indicates the unit sphere.

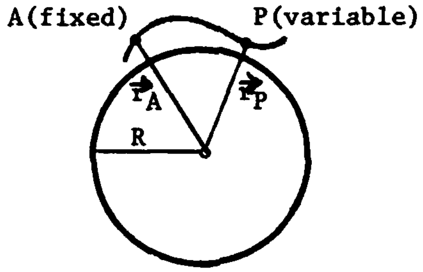


Figure 2

The first two terms of eq. (16), involving the anomaly $\Delta\hat{g}$, represent the "signal" components of ξ and η. The following three terms constitute computable topographic "noise". Permitting random-type errors r_ξ and r_η, eq. (46) can be written in analogy with eqs. (40) and (41) as

$$\begin{bmatrix} \xi \\ \eta \end{bmatrix} = \begin{bmatrix} \hat{\xi} \\ \hat{\eta} \end{bmatrix} + \begin{bmatrix} \delta_{\xi_t} \\ \delta_{\eta_t} \end{bmatrix} + \begin{bmatrix} r_\xi \\ r_\eta \end{bmatrix} \qquad (47)$$

The numerical determination of the three topographic terms of eq. (46) is a complex task, which can, however, be accomplished without inherent difficulties by means of high-speed computers. The estimation of ξ and η from a set of given vertical deflections can be achieved by means of spatial covariance functions. It is clear from the above analysis that the optimal estimation method outlined in section 4 requires the simultaneous consideration of eq. (47). Otherwise, the linear aggregates involving ξ and η in eqs. (35) and (36) are to be estimated with some error from a limited set of ξ and η in order to obtain solutions without linear aggregates. Repetitive surveys would reduce the effects of instrument errors. In summary, the determination of deflections of the vertical in strongly mountainous terrain requires a much increased computational effort to obtain minimum error accuracies. Without such effort, there is a degradation in accuracy.

6. UTILIZATION OF GENERATED GRAVITY VECTOR COMPONENT INFORMATION.

Rapid economical gravity vector component determination is necessary, useful, or potentially useful for the following purposes:

(a) Transformation of astronomic coordinates and astronomic azimuths into ideal terrestrial-geodetic coordinates and geodetic azimuths, requiring knowledge of absolute deflections of the vertical, indispensible for long range ballistic missiles.

(b) Establishment of a three-dimensional anomalous gravity potential T or analytical upward continuation of first order derivatives of Δg, η, and ξ from discrete measurements thereof at the earth's surface, to extend over relatively large areas, indispensible for long range ballistic missiles with inertial guidance.

(c) Adjustment of geodetic networks.

(d) Geophysical prospecting including estimation of the derivative $\frac{\partial \Delta g}{\partial r} = 2GR^{-1}N + GR^{-1}\xi \operatorname{tn} \phi - G \left(\frac{\partial \xi}{\partial x} + \frac{\partial \eta}{\partial y}\right)$ in semi-flat terrain where N stands for geoidal undulation or height anomaly.[4]

(e) Determination of geoidal differences on land.

(f) Construction of regional gravity anomaly and vertical deflection networks with mesh sizes $\Delta x = \Delta y$ of about 5 Km as input information for highly accurate gravity programmed inertial surveying systems.

(g) Accuracy improvement of medium range ballistic missiles in conjunction with meteorological corrections and of cruise missiles by utilization of vertical deflection data in accessible areas.

7. CONCLUSIONS. The determination of vertical gravity vector components or gravity anomalies by means of inertial and boundary data has been very promising as to rapidity and accuracy. High accuracy estimation of horizontal gravity vector components or deflections of the vertical from inertial and boundary data requires the installation of gryoscopes, accelerometers, and velocity quantizers with small error variances and short correlation times, the development of sophisticated estimation methods consistent with an outlined structure, and, probably, a three-dimensional Kalman filter. Deflection accuracy in strongly mountainous terrain is degraded. For this reason, a maximum rms error requirement of about 0.3 arcsec in both semi-flat and mountainous terrain may only be achieved by a gradiometer-aided inertial system. Such

---

[4] See Telford et al [11], pp. 7-104, for geophysical applications and Heiskanen and Moritz [12], p. 117, for the equation.

a system would also be beneficial for the location of smaller subterraneous mass anomalies. An effective estimation of gravity anomalies and deflections of the vertical in mountainous terrain from data at points separated by distances of the order 5-10 Km is possible with considerable computational effort under utilization of the developed equations and employment of spatial covariance functions. Numerous military and non-military applications of gravity vector component information generated by inertial and auxiliary data do exist. Gradiometer-aided inertial systems offer additional advantages.

REFERENCES

[1]. Litton Systems, Inc. Geodetic Platform Development Program, Public. No. 12987, Woodland Hills, CA 91364, 1974.

[2]. Huddle, J. The Measurement of the Change in the Deflection of the Vertical betwen Astronomic Stations with a Schuler-Tuned Inertial System. Litton Guidance & Control Systems, Woodland Hills, CA 91364, 1977.

[3]. Litton Systems, Inc. P-1000 Inertial Instrument, A-1000 Accelerometer. Public. No. 12391 A. Woodland Hills, CA 91364, 1973.

[4]. Litton Systems, Inc. Addendum to Feasibility Test Program for Measurement of Gravity Anomaly Changes Using 2 Micro-G Accelerometer in Inertial Platform. Document No. 402414. Woodland Hills, CA 91364, 1975.

[5]. Litton Systems, Inc. Proposal for Position and Azimuth Determining System (PADS), Vol. II - Technical Proposal. Woodland Hills, CA 91364, Sept. 1970.

[6]. Moritz, H. Current Determination of Vertical Deflections. Proc. 1st Internatl. Symp. on Inertial Technology for Surveying and Geodesy, Ottawa, Canada, Oct. 1977.

[7]. Schwarz, K. P. Inertial Surveying Systems - Experience and Prognosis. Paper, presented at the FIG-Symposium on Modern Technology for Cadastre and Land Information Systems, Ottawa, Canada, Oct. 1979.

[8]. Huddle, J. The Measurement of the Change in the Deflection of the Vertical with a Schuler-Tuned North-Slaved Inertial System. Litton Guidance & Control Systems, Woodland Hills, CA. 91364, 1977.

[9]. Baussus von Luetzow, H. A New Method for the Determination of Deflections of the Vertical from Astrogeodetic and Inertially Derived

Data. Paper, presented at the Spring Meeting of the Am. Geoph. Union, Baltimore, MD., May 1981.

[10]. Baussus von Luetzow, H. On the Interpolation of Gravity Anomalies and Deflections of the Vertical in Mountainous Terrain. Proc. VIth Internatl. Symp. on Geodetic Computations, Munich, W. Germany, Aug. - Sept. 1981.

[11]. Telford, W., L. Geldert, R. Sheriff, and D. Keys, Applied Geophysics, Cambridge Univ. Press, Cambridge CB2 IRP, England, 1976.

[12]. Heiskanen, W. and H. Moritz. Physical Geodesy, W. H. Freeman and Co., San Francisco and London, 1967.

MULTIPLE LAUNCH ROCKET SYSTEM (MLRS) FUZE (U)

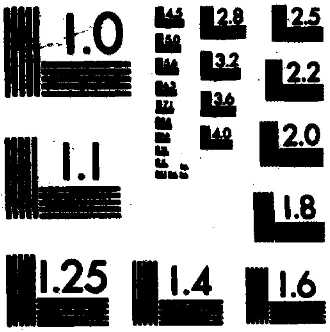
JOHN J. BERTIN, DR.
THE UNIVERSITY OF TEXAS AT AUSTIN
AUSTIN, TEXAS 78712
*RICHARD L. GOODYEAR, MAJ
U.S. MILITARY ACADEMY
WEST POINT, NEW YORK 10996

1. INTRODUCTION

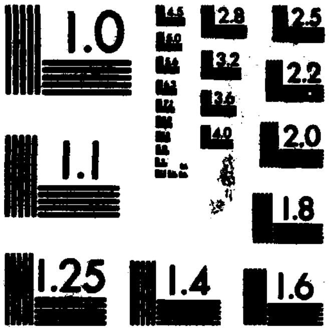
When the U.S. Army began limited production of the fuze for the rocket warhead of the Multiple Launch Rocket System (MLRS), it was necessary to develop improved methods of testing fuzes as they came off the assembly line. Since the fuze contained a fluidic power supply that had never been type-classified, methods associated with previous battery-powered fuzes were inadequate for the production specification. Also during the validation and maturation phases of the development program, considerable testing in the laboratory, in wind tunnels, and in flight tests showed that laboratory tests in early specifications did not indicate adequately how a fuze would perform in actual flight. A key deficiency was the inability to duplicate in the laboratory the conditions near the summit of the rocket's trajectory. Operating the fluidic generator from an air source up to 14 kilopascals above atmospheric pressure approximated the expected pressure difference between the entrance and exit of the generator, but it failed to duplicate the effects of the air's speed.

The power supply project team at the U.S. Army Harry Diamond Laboratories (HDL) proposed a solution in the form of a relatively inexpensive high altitude chamber, which would duplicate both of these flight variables, and which could be incorporated into the fuze contractor's test equipment. Implementation of this proposal required a better understanding of the conditions of the air flow around the fuze ogive and through the generator. Consequently, HDL conducted a test program in the Aerodynamic Wind Tunnel at the Arnold Engineering Development Center. Test conditions included free-stream Mach numbers from 0.95 to 1.3 over a range of free-stream densities that simulated altitudes from 15.2 Km to 21 Km. The fuze ogive was instrumented with both static and pitot pressure probes, from which the pressure data were analyzed to describe the internal and external flow fields.

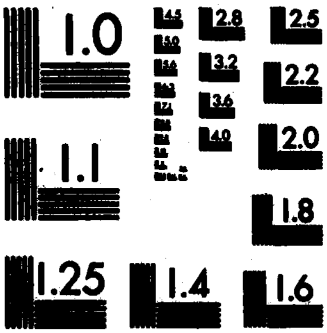
AD-A120 811 PROCEEDINGS OF THE 1982 ARMY SCIENCE CONFERENCE HELD AT 2/6
THE UNITED STATES..(U) DEPUTY CHIEF OF STAFF FOR
RESEARCH DEVELOPMENT AND ACQUISITIO.. 18 JUN 82
UNCLASSIFIED F/G 5/2 NL


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

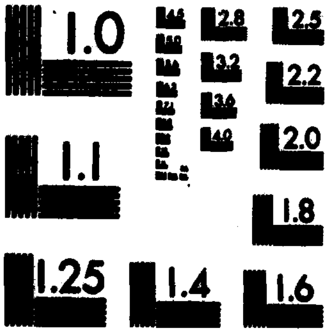
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

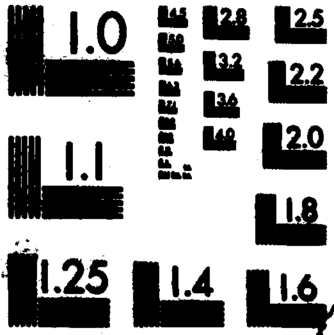
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

2. EXPERIMENTAL PROGRAM

The general features of the nose cone are illustrated in the sketch presented in Figure 1. A sectioned drawing has been used to present the details of the external geometry and of the generator installation. The air that drives the fluidic generator enters the "inlet hole" in the stagnation region of the blunt face, passes through the generator assembly and into the (internal) chamber. It then flows through the exhaust ports into the (external) cavity. This use of the words "chamber" for the internal region and "cavity" for the depression in the external surface will be retained throughout this report. The 24 exhaust ports, which are 0.269 cm in diameter, are equally distributed circumferentially and are located in a shallow cavity that begins 3.259 cm ($x = 0.423\, x_o$) aft of the tip of the fuze. As shown in figure 1, x is the longitudinal position, as measured from the model nose. The reference length, x_o, was chosen to be 7.696 cm, the coordinate of the most-downstream pressure tap.

The nose cone used in the initial phase of the test program was instrumented with 29 surface pressure taps (or orifices), of which 21

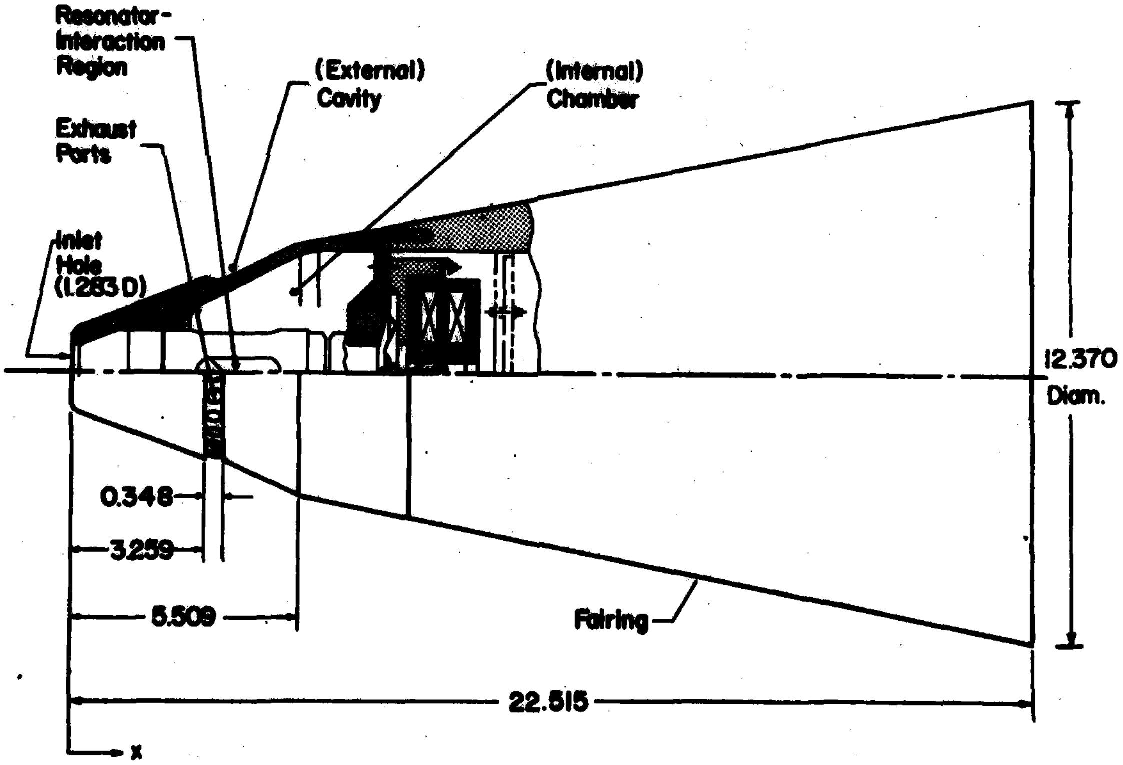

Resonator-
Interaction
Region
(External)
Cavity
(Internal)
Chamber
Exhaust
Ports
Inlet
Hole
(1.283 D)
12.370
Diam.
0.348
3.259
5.509
Fairing
22.515
x

Figure 1. - Sketch of the Army/MLRS Fluidic-Generator Fuze-Assembly.

were on the external surface and 8 were on the internal surface. The locations of the orifices for which data are presented in this paper are illustrated in Figure 2.

3. RESULTS

Static pressures and Pitot pressures were measured both on the external surface and on the internal surface of a full-scale fluidic generator/fuze assembly of the MLRS in the Aerodynamic Wind Tunnel (4T) at AEDC. The pressure data from these tests are discussed in this section.

3.1 External Pressures for 0° Angle of Attack

3.1.1 Pressures on Flat Face (x = 0.0)

A strong, detached shock wave forms in front of a blunt body when it is placed in a supersonic stream. The shape of the bow shock wave

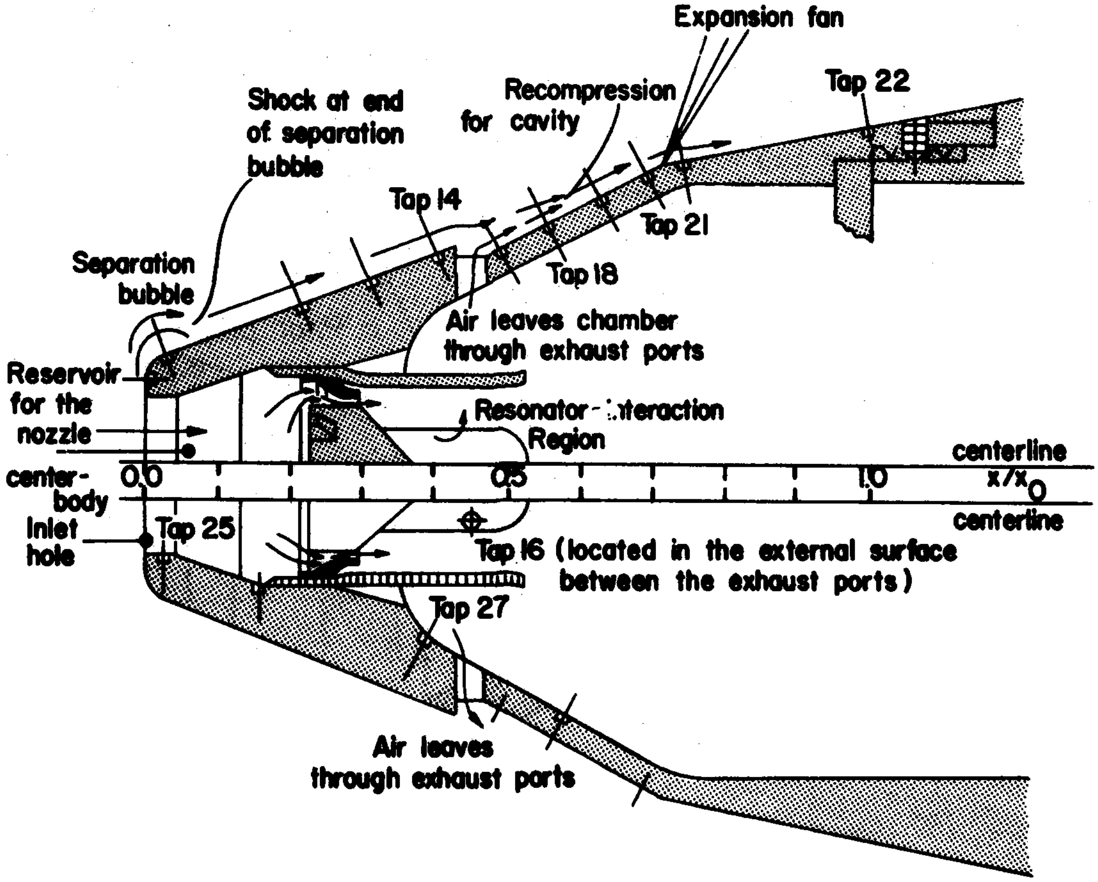

Expansion fan
Tap 22
Shock at end of separation bubble
Recompression for cavity
Tap 14
Tap 21
Separation bubble
Tap 18
Air leaves chamber through exhaust ports
Reservoir for the nozzle
Resonator-interaction Region
centerline
x/x_0
center-body
0.0
0.5
1.0
Inlet hole
Tap 25
centerline
Tap 16 (located in the external surface between the exhaust ports)
Tap 27
Air leaves through exhaust ports

Figure 2. - Sketch illustrating instrumentation locations and features of proposed flow model.

for supersonic flow past the nose cone is shown in the schlieren photograph from a previous test program (1) that is presented in Figure 3. Since the Mach number of the flow represented by the schlieren photograph is greater than those of the present program, the stand-off distance of the bow shock wave is much less than would be the case for the present flows. This is to be expected. Nevertheless, the photograph provides valuable insights into the flow. Because the bow shock wave is curved, the static pressure on the surface should decrease from a value of the stagnation pressure downstream of a normal shock wave, p_{t2} , at the axis of symmetry to the sonic value at the corner. This pressure variation is illustrated by the pressure distribution that was presented elsewhere (2) for a free-stream Mach

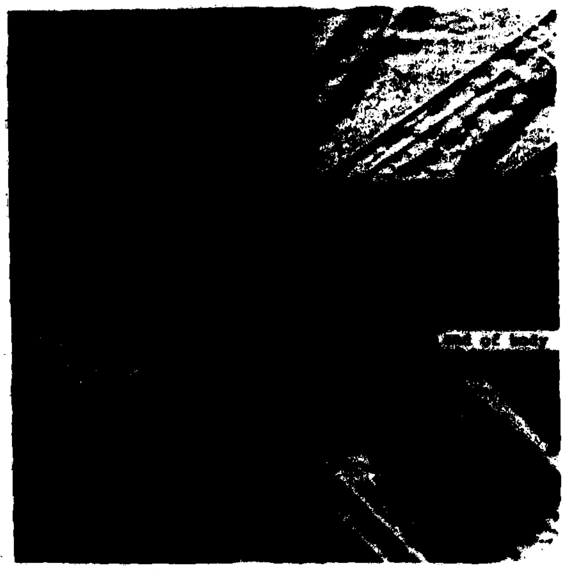


Figure 3. - Schlieren photograph from the tests of ref. 1, $M_\infty = 1.5$.

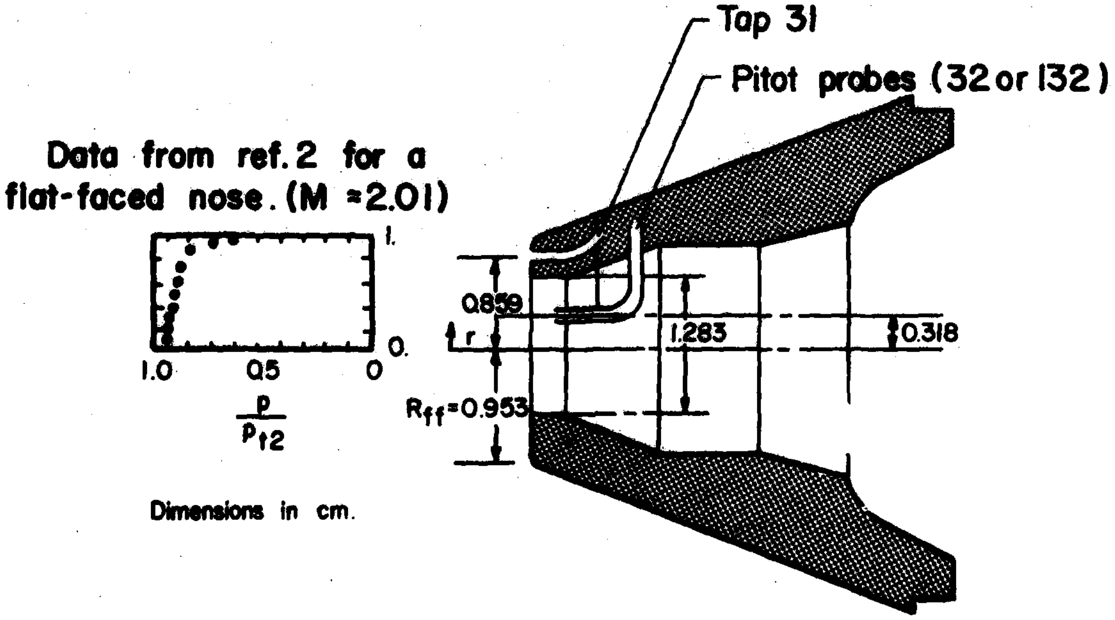

Tap 31
Pitot probes (32 or 132)
Data from ref. 2 for a flat-faced nose. (M =2.01)
1.
0.
1.0
0.5
0
$\frac{p}{p_{t2}}$
0.859
r
1.283
0.318
R_{ff}=0.953
Dimensions in cm.

Figure 4. - Nose-region instrumentation and pressure distributions for a flat-faced cylinder (ref. 2).

number, M_∞, of 2.01 (see Figure 4). Although the presence of the inlet hole will modify the flow on the face of the fuze, the pressure distributions for the present tests will be similar to that of Figure 4 despite the difference in Mach number.

Included with the pressure distributions in Figure 4 are the pressure-instrumentation locations for two model variations tested. For selected tests a pitot probe was positioned in the inlet hole. There were two static-pressure orifices on the flat face of the nose cone, between the inlet hole and the corner radius. They were approximately 0.859 cm from the axis of symmetry, i.e., $r = 0.90 R_{ff}$. The experimentally determined pressure coefficients, C_p, for orifice 31 are presented in Figure 5 at density altitudes (AD) of 18,290 m and 21,030 m. Included are the experimental values of C_p for the two total-pressure probes that were 0.318 cm from the axis, in the inlet hole. These experimentally determined C_p are compared with the theoretical C_p, as calculated for the flow downstream of a normal shock wave, using the relation:

$$C_{p,t2} = \left(\frac{p_{t2}}{p_\infty} - 1\right)\frac{2}{\gamma M_\infty^2}, \qquad (1)$$

where $C_{p,t2}$ is the stagnation pressure coefficient, p_∞ is free-stream static pressure, and γ is the ratio of specific heats. The ratio p_{t2}/p_∞ was determined as a function of M_∞ using normal shock wave tables (3).

The experimental values of C_p determined using the Pitot probes in the inlet hole are in very good agreement with the theoretical values. This is to be expected, since the probes are near the axis of symmetry (where the bow shock wave is most nearly normal) and, being Pitot probes, they provide a measure of the stagnation pressure. The surface static pressures at tap 31 are below the theoretical predictions. This is also to be expected. When a flat-faced cylinder with a "slender" conical afterbody is exposed to a supersonic stream, the bow shock wave is curved and the sonic "line" is fixed at the corner of the flat-

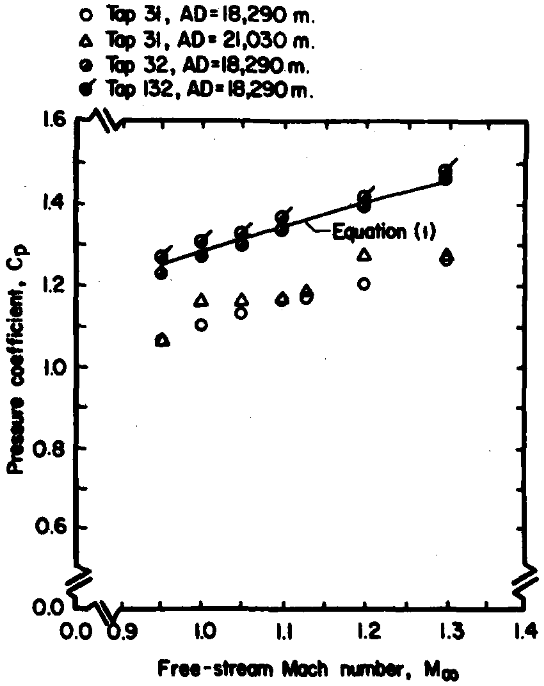


Figure 5. - Pressure coefficients for flat face as a function of free-stream Mach number.

faced portion of the nose cone. As a result, the static pressure decreases with r over the flat face from the "stagnation pressure" at the axis of symmetry to the sonic value at the corner. Using the data presented in Figure 4 (which are for $M_\infty = 2.01$), one would expect the local static pressure at $r = 0.90\ R_{ff}$ to be approximately $0.9\ p_{t2}$. As a result, the C_p for this orifice would be approximately 1.22 when $M_\infty = 1.3$. This value is in reasonable agreement with the data of Figure 5, considering the difference in the free-stream Mach numbers and the differences in the flow field due to the presence of the inlet hole.

3.1.2. Pressures on Conical Surface of Ogive

Additional information about the flow field can be determined from the schlieren photograph in Figure 3. Of interest is the flow over the first (forward) conical segment. A rapid expansion of the flow just past the corner is terminated by a shock wave near $x = 0.1\ x_0$. It is assumed that the shock wave results when the flow external to a separation bubble (which exists at the corner) turns when it reattaches to the surface. Furthermore, although the cavity (containing the exhaust ports) is relatively long, the flow appears to "jump over" the cavity; i.e., the cavity is open, since there is no recompression shock wave near the end of the cavity. However, the pressure data from the present tests indicate that the flow does not completely clear the cavity since the pressures indicate a (weak) recompression of the flow near the midpoint of the cavity.

The experimentally-determined C_p distributions for $M_\infty = 1.3$ at a density altitude of 18,290 m are presented in Figure 6. Included in Figure 6 is an outline of the configuration, showing pertinent geometric features. The detailed sketch of the geometry of the nose cone, Figure 2, illustrates the locations of the pressure taps and some features of a proposed model for the external flow. Although the flow expands rapidly around the corner, it cannot accelerate fast enough to remain attached to the surface, and a separation bubble forms at the corner. This assumption is supported by the pressure measured at the orifice in the "separated region" (at $x = 0.030\ x_0$) which is relatively low.

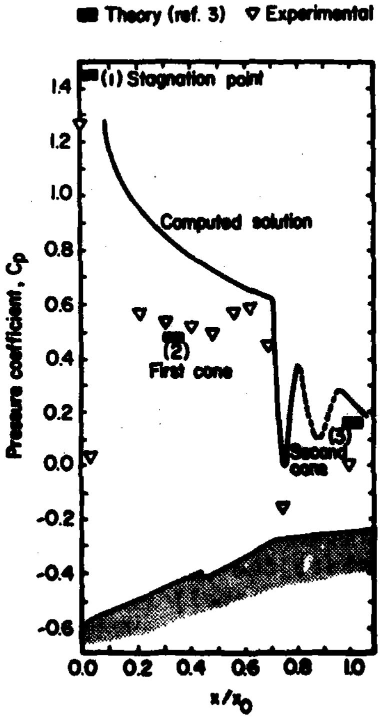


Figure 6. - Comparison between theoretical and experimental pressure distributions, $M_\infty = 1.3$, AD = 18,290 m.

Downstream of the separation bubble, the flow over the external surface of the nose cone was calculated using two different approaches. Theoretical values of C_p were calculated using standard normal shock-wave tables and graphs for sharp cones. Values are presented

(1) for the stagnation point ($C_p = 1.45$),
(2) for the forward conical segment ($C_p = 0.48$, the value for a sharp cone whose half-angle is 20.5°), and
(3) for the second conical segment ($C_p = 0.16$, the value for a sharp cone whose half-angle is 10.75°).

Also included is the theoretical distribution computed using the University of Texas' Missile Aerodynamics Code, which is similar to that described by Moore and Swanson (4). In preparing the geometric input for the computer solution, the presence of the cavity in the external contour was omitted and the nose cone was represented by two conical frustums. Because of the extreme bluntness of the nose cone, the computed solution should be considered approximate.

Both the computed solution and the "sharp-cone" approximations (3) provide a reasonable estimate of the pressures on the downstream (second) conical segment, i.e., for $x \geq 0.716\ x_o$. The computer solution even predicts the rapid "overexpansion" that takes place downstream of the intersection of the two conical segments. An oscillatory character which appeared in the computed solution for the pressure distribution for the second conical segment is represented in Figure 6 by a broken line. The broken-line representation was used since the variation may be a computational peculiarity and not a flow phenomenon.

The most significant differences between the theoretical values and the experimental values occur on the forward conical surface and in the cavity in this region. Note that the upstream edge of the cavity is at $x = 0.423\ x_o$. The experimental pressures at the first tap of the cavity (tap 17 at $x = 0.482\ x_o$) are essentially equal to those upstream of the cavity. The fact that the pressure did not decrease when the flow separated at the cavity is attributed to the relatively shallow depth of the cavity. The pressure increase at the three pressure taps in the downstream region of the cavity ($0.551\ x_o \leq x \leq 0.689\ x_o$), reflects the recompression of the flow as it reattaches to the surface. Thus, if one assumes that a tangent-cone approximation is valid for the forward conical surface and that a recompression pressure rise occurs at the downstream end of the cavity, one can explain the fact that the pressures measured at the taps between $0.551\ x_o$ and $0.689\ x_o$ are above the Ames theoretical estimates (3). Since both theoretical models are approximate, close correlation with the data should not be expected. Each assumed model probably contains elements of the actual flow field.

The effect of M_∞ on the pressure distribution is illustrated by the data presented in Figure 7. This pressure distribution is for a density altitude of 21,030 m. There is a consistent trend with Mach number at all the orifices from $x = 0.030\ x_o$ to $x = 1.000\ x_o$. That is, the ratio p/p_{t2}

at a particular orifice decreases as the Mach number increases. The variation is greatest at the two extreme orifices. The Mach-number dependent variation at $x = 0.030\ x_o$ indicates a change in the characteristics of the separation bubble for these transonic flows. At the last orifice, the pressure ratio varies by almost a factor of two. However, the experimental values of p/p_{t2} at this orifice vary relatively little for $1.1 \leq M_\infty \leq 1.3$. Thus, the spread in the data for the three lowest Mach numbers suggests that a significant change occurs in the downstream region of the flow field for these transonic flows. But these transonic changes in the flow field occur far enough downstream (on the second conical segment) that they do not affect the flow through the fluidic generator.

To establish if this Mach-number dependence has an analytical basis, let us examine the C_p as given by the small deflection approximation for supersonic flow,

$$C_p = \left(\frac{p}{p_\infty} - 1\right)\frac{2}{\gamma M_\infty^2} = \frac{C_1}{\sqrt{M_\infty^2 - 1}} , \qquad (2)$$

where C_1 is a constant (which is dependent on M_∞ and the local flow direction). The equation can be rearranged to yield:

$$\frac{p}{p_\infty} = \frac{\gamma C_1 M_\infty^2}{2\sqrt{M_\infty^2 - 1}} + 1 . \qquad (3)$$

Rewriting this expression in terms of the pressure ratio used in Figure 7, we obtain

$$\frac{p}{p_{t2}} = \frac{p_\infty}{p_{t2}}\left(\frac{0.7C_1 M_\infty^2}{\sqrt{M_\infty^2 - 1}} + 1\right). \qquad (4)$$

Using the Ames tables (3) to obtain the value of p_∞/p_{t2} as a function of Mach number, one finds the following values.

$$\begin{aligned} \frac{p}{p_{t2}} &= 0.5595\ [0.7C_1(2.8903)+ 1] \text{ for } M_\infty = 0.95 \\ &= 0.4689\ [0.7C_1(2.6402)+ 1] \text{ for } M_\infty = 1.10 \\ &= 0.3685\ [0.7C_1(2.0345)+ 1] \text{ for } M_\infty = 1.30. \end{aligned}$$

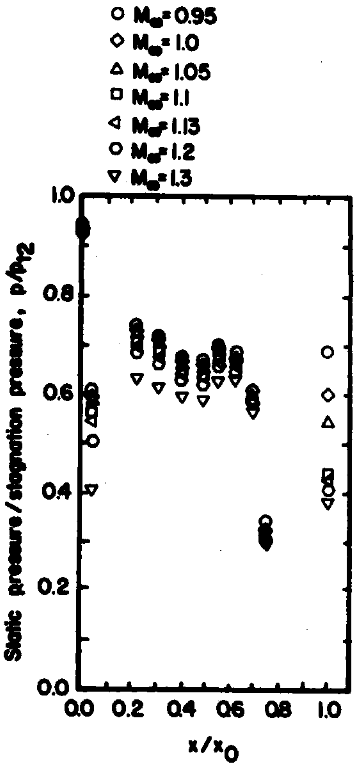


Figure 7. - Effect of Mach number on pressure distribution for AD = 21,030 m.

Using this relatively simple approximation, one would expect the pressure ratio p/p_{t2} to decrease with Mach number for these test conditions. Thus, although one would predict a greater decrease in the pressure ratio using equation (4) than is actually exhibited by the data, the small-deflection approximation provides a qualitative substantiation of the observed Mach-number dependence.

Although the pressures measured on the conical surface exhibit an inverse dependence on M_∞, the pressure ratio (p/p_{t2}) for locations on the blunt face should be independent of the Mach number for supersonic flows. The data from the orifice on the flat face (see the pressure measurements for x = 0.0, presented in Figure 7) exhibit this Mach-number independence.

3.2 Flow Through Fluidic Generator and Internal Pressure Measurements for 0° Angle of Attack

As has been discussed, the air which activates the fluidic generator enters the inlet hole in the stagnation region, passes through the resonator-interaction region of the generator assembly and then into the internal chamber, before it finally flows through the exhaust ports. A sketch of this internal flow pattern and of the relevant pressure instrumentation is presented in Figure 8. The mass-flow rate through the fluidic generator was calculated with the pressures measured at the orifices of Figure 8, using two different assumptions regarding the air flow through the fluidic generator.

Flow Model 1.--Because the area of the annular gap around the nozzle centerbody is so much less than the area of the inlet hole, the air velocity forward of the centerbody is relatively small. Thus, this region serves as a reservoir (or stagnation chamber) for the nozzle centerbody. The static pressures both at tap 25 ($x = 0.026\ x_o$) and at tap 26 ($x = 0.132\ x_o$) are essentially equal to p_{t2} over the entire range of Mach number and of density altitude.

It is assumed that the flow accelerates isentropically from this "reservoir" through the annular gap, reaching the speed of sound at the minimum cross-sectional area of the annular gap. Thus, the centerbody serves as a throat, "choking" the flow over the entire range of test conditions for this program. Fliegner's formula (5) for the choked flow of perfect air along with the minimum cross-sectional area of the annulus can be used to calculate the mass-flow rate, $\dot{m}_1$:

$$\frac{\dot{m}_1}{\rho_\infty U_\infty A_{In}} = \frac{0.0386[p(25)]}{\rho_\infty U_\infty \sqrt{T_t}} , \qquad (5)$$

where p(25) is the pressure at tap 25 (see Figure 2), ρ_∞ = free-stream density (kg/m^3), U_∞ = free-stream velocity (m/s), and T_t = stagnation temperature (K). The local mass-flow rate has been divided by $\rho_\infty U_\infty A_{In}$

(which is equal to the free-stream mass-flow rate across an area equal to that of the inlet hole) in order to obtain a dimensionless parameter, which is independent of the Mach number and of the density altitude over the range of test conditions.

Because the air must go around the corner of a flat-faced cylinder as it enters the nozzle centerbody (producing a separation bubble) and because of the presence of boundary layers on both surfaces of a narrow annular gap, it is expected that the effective throat area is significantly less than the minimum cross section of the annular gap (which is the area used in eq. (5)). Thus, it is expected that the non-dimensionalized values of $\dot{m}_1$, as calculated using equation (5) (which are represented by the half-filled symbols of Figure 9), overestimate the actual mass-flow rate.

It is believed that the flow is indeed choked by the nozzle centerbody, even though the static pressures in the (downstream) internal chamber are approximately 0.7 p_{t2} . The reason for the chamber value of 0.7 p_{t2} follows. The pressure in the internal chamber is governed (1) by the pressure in the external cavity (which is established by the external flow field), since (as will be discussed) the flow through the exhaust ports is not choked, and (2) by the pressure drop across the exhaust ports, which is governed by the mass-flow rate, as shown in equation (6a). The mass-flow rate is established by the choking of the flow through the

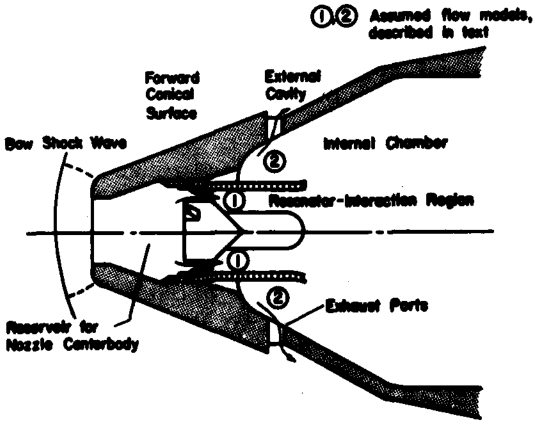

①,② Assumed flow models, described in text
Forward Conical Surface
External Cavity
Bow Shock Wave
Internal Chamber
②
①
Resonator-Interaction Region
①
②
Exhaust Ports
Reservoir for Nozzle Centerbody

Figure 8. - Nomenclature for flow models.

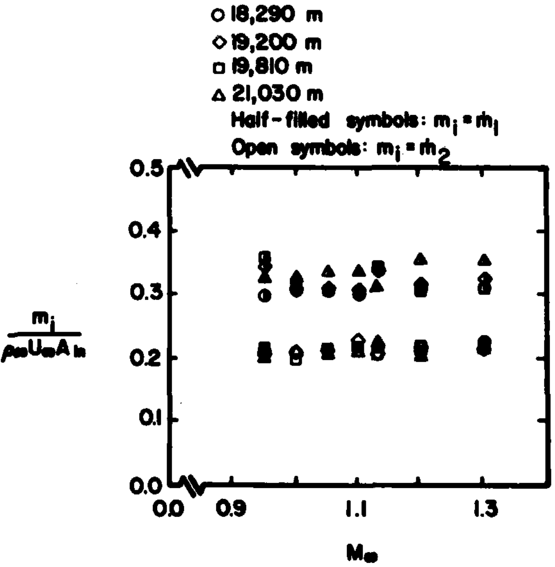


Figure 9. - Mass-flow rate through fluidic generators.

nozzle centerbody. The static pressure increases from the sonic value (0.53 p_{t2}) at the throat of the nozzle centerbody to 0.7 p_{t2} , the value in the internal chamber, as it flows through the resonator-interaction region of the generator assembly (see Figures 1 and 8) and into the internal chamber.

The conclusion that the flow is choked not at the exhaust ports but at the nozzle centerbody over the range of test conditions is very important. It has been shown that the pressure ratio (p/p_{t2}) for the separated region of the external cavity is a well-behaved function of the free-stream Mach number and is independent of density altitude. Since the pressure drop across the exhaust ports is relatively small and is a function of the mass-flow rate through the fluidic generator, the pressure in the internal chamber (which is the "back pressure" for the resonator-interaction region of the fluidic generator) follows the external pressure. However, since the flow is choked by the nozzle centerbody, the inlet pressure for the resonator-interaction region is a fixed fraction of p_{t2} (i.e., 0.53, the sonic value). It is expected that the details of the flow field in the resonator depend on the pressure differences across it.

Flow Model 2.--For the second flow model, the volumetric flow-rate, Q, was calculated using the equation for subsonic flow through a sharp-edged, circular orifice (6):

$$Q = 0.6YA\sqrt{\frac{2\Delta p}{\rho}} \quad , \tag{6a}$$

where for the orifice ratios and for the Reynolds numbers of the present tests,

$$Y = \left\{1.0 - 0.3\ \frac{\Delta p}{p_1}\right\} \tag{6b}$$

Furthermore, it is assumed that the acceleration of the flow from the internal chamber through the exhaust ports is such that

$$\rho \propto p \tag{6c}$$

Combining these relations, one can calculate the mass-flow rate for this model using the relation

$$\frac{\dot{m}_2}{\rho_\infty U_\infty A_{in}} = \frac{0.158\ p(16)}{\rho_\infty U_\infty \sqrt{T_t}} \left\{0.7 + 0.3\,\frac{p(16)}{p(27)}\right\} \sqrt{1.0 - \frac{p(16)}{p(27)}} \tag{7}$$

The mass-flow rates thus calculated are presented as a function of M_∞ in Figure 9 (as the open symbols). The experimentally determined values of the pressure ratio [p(16)]/[p(27)] indicate that the flow in this region is incompressible. Thus, the assumptions made in developing flow model 2 are quite realistic. Since the two pressures needed to calculate the flow field were measured directly, it is believed that the mass-flow rates calculated using equation (7), i.e., $\dot{m}_2$, are the more realistic of the two flow models.

4. CONCLUSIONS

This analysis showed that pressures measured at the inlet hole to the fluidic generator are in close agreement with those given by normal shock relations. Over the external surface of the ogive, the ratio of local static pressure to stagnation pressure measured at the nose, p/p_{t2} decreased as the free stream Mach numbers increased. Within this trend existed regions of expansion and compression which were a result of the ogive's geometry and the injection of the air exhausted from the fluidic generator. Theoretical models predicted reasonably well the dependence on Mach numbers of p/p_{t2} as well as the relative values of p/p_{t2} at different points on the ogive. Pressures measured at the entrance and exit of the region occupied by the generator permitted calculation of the mass flow rate through the fuze. This analysis verified that the key factors that should be duplicated in laboratory tests were the pressures at the generator's entrance and at the exhaust ports in the ogive; but the most significant result was the accumulation of a data base of the actual pressures to be associated with desired combinations of Mach number and altitude. Consequently, the high altitude chamber proposed to test MLRS fuzes has been incorporated into the product specification and used with confidence in the quality assurance program of the contractor.

LITERATURE CITED

1. R. L. Goodyear and H. Lee, Performance of the Fluidic Power Supply for the XM-445 Fuze in Supersonic Wind Tunnels, HDL-TM-81-4, Feb. 1981, Harry Diamond Laboratories.
2. J. C. Boison and H. A. Curtiss, An Experimental Investigation of Blunt Body Stagnation Point Velocity Gradient, ARS Journal, Feb. 1959, Vol. 29, No. 2, pp. 130-135.
3. Ames Research Staff, Equations, Tables, and Charts for Compressible Flow, Report 1135, 1953, NACA.
4. F. G. Moore and R. C. Swanson, Jr., Aerodynamics of Tactical Weapons to Mach Number 3 and Angle of Attack 15°, Part I--Theory and Application, NSWC/DL TR-3584, Feb. 1977, Naval Surface Weapons Center.
5. A. H. Shapiro, The Dynamics and Thermodynamics of Compressible Fluid Flow, Ronald Press, New York, 1953.
6. Staff, Flow of Fluid Through Valves, Fittings, and Pipes, Technical Paper No. 410, The Crane Company.

THERMAL FATIGUE BEHAVIOR OF FP ALUMINA/MAGNESIUM COMPOSITES

*Ramakrishna T. Bhatt, PhD
Propulsion Laboratory, USARTL(AVRADCOM)
Cleveland, Ohio

Introduction

Composites of magnesium alloys reinforced with FP-Alumina are being considered for aerospace applications because of their high specific strength and modulus, and because of their relative ease of fabrication. Some of these applications require repeated high temperature exposure of the composite material for extended periods of time. In a recent study, Bhatt, et al.(1) have evaluated the effects of isothermal and cyclic exposure on the room temperature mechanical properties of unidirectionally reinforced FP-Al_2O_3/EZ33 magnesium composites. Results of this study indicate no significant loss in the room temperature axial tensile strength and dynamic flexural modulus of composites thermally cycled between 50° and 250°C or of composites isothermally heated at 350°C for up to 150 hours from the strength and modulus data for the as-fabricated composites. In contrast, thermal cycling between 50°and 300°C caused considerable loss in both room temperature strength and modulus. The major causes of strength degradation were attributed to softening and cracking of the matrix and debonding of the fibers from the matrix.

The objective of this investigation was to study the effects of matrix composition on thermal fatigue behavior of FP-Al_2O_3/magnesium composites as determined by tensile stength and modulus measurements after cyclic exposure.

In this study, 55 volume percent fiber unidirectionally aligned composites were exposed in air to temperatures of either 250° or 350°C for up to 3000 cycles. Isothermal exposure at 350°C was used as a baseline to evaluate the additional effect of cycling. Metallographic and fractographic studies were made on untreated and thermally cycled composite material in an attempt toward understanding the failure mechanisms involved.

Experimental

The FP-Al_2O_3/QH21A Mg composites used in this study were fabricated by the Dupont Pioneering Research Laboratory using molten metal infiltration techniques. The nominal FP-Al_2O_3 fiber content used was 55 volume percent. The fibers were aligned unidirectionally in a QH21A magnesium alloy matrix having a composition of 2 to 3% Ag, 0.6 to 1% Th, 0.6 to 1.5% rare earths, 0.4 to 1% Zr, and bal. Mg. Plates of 0.25 cm thickness were cast and cut into specimens 12.7 cm long and 1.25 cm wide. The fiber orientation was either parallel to the specimen length (0°) for axial testing or perpendicular to the specimen length (90°) for transverse testing. Thermal cycling was done by alternately dipping a frame supporting six specimens into a hot (250° or 350°C) fluidized sand bath and then into a cold bath that equilibrated near 50°C. Each complete thermal cycle lasted for six minutes. Typical time-temperature profiles of composite specimens cycled to 250° or 350°C are shown in Figure 1. The time at temperature during each cycle was approximately two minutes.

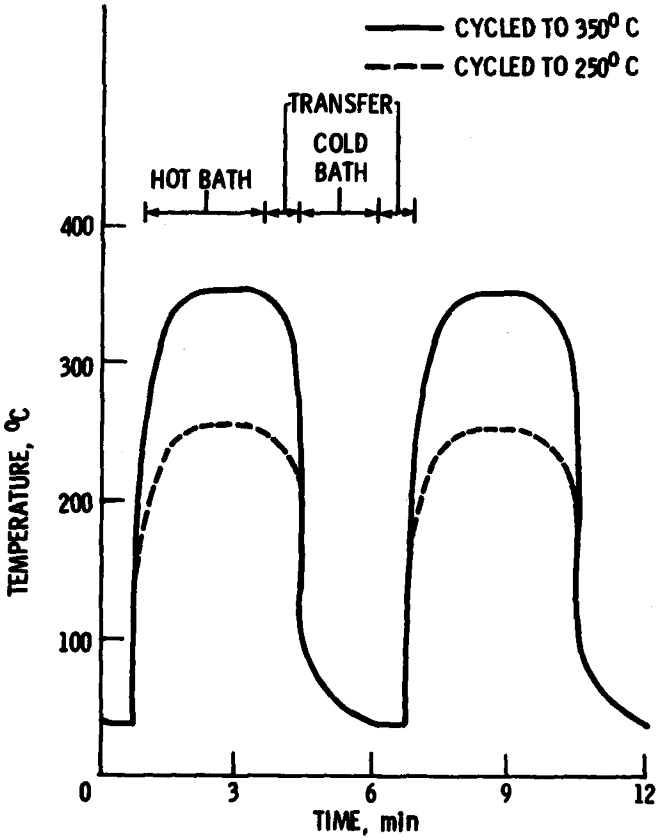


Figure 1 Typical thermal cycle temperature-time profiles.

Similar composite specimens were also isothermally heated at 350°C in a sand bath for periods up to 100 hours, which corresponds to a time equivalent to the time at temperature for 3000 cycle experiments.

After cycling to a predetermined number of cycles or isothermal heating to set time periods, the specimens were removed from the bath and aluminum doublers were adhesively bonded to the specimen ends.

Tensile testing was done in an Instron testing machine equipped with wedge-type grips. The specimens were pulled to failure at a constant cross head speed of 0.126 cm/min.

A flexural modulus test was used for measuring the dynamic modulus of the composite specimens. The test method and the equations used for calculating dynamic modulus of the composites are described in Reference 1.

Results and Discussions

The room temperature flexural moduli of the 55 fiber volume percent FP-Al_2O_3/QH21A Mg composites cycled to 250° or 350°C to a maximum of 3000 cycles are shown in Figure 2. The flexural moduli of untreated composites are also

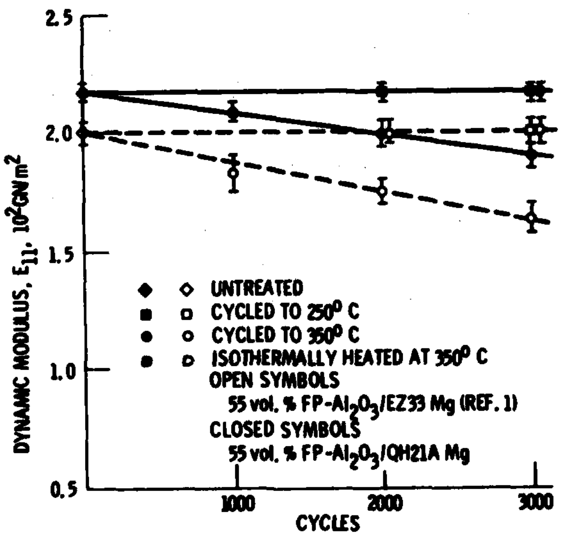


Figure 2 Room temperature dynamic moduli of FP-Al_2O_3/magnesium composites after cycling to 250°C or 350°C for indicated number of cycles. Dynamic moduli of composites isothermally heated at 350°C for 100 hours are shown. Dynamic moduli of untreated composites are shown at zero cycles.

shown in Figure 2 for baseline comparison. The data points represent the range and average value for at least three determinations. The averaged data of Bhatt, et al. (1) for 55 fiber volume percent $FPAl_2O_3/EZ33$ Mg composites thermally cycled to 250° or 350°C for up to 3000 cycles are shown in this figure for comparison. The modulus values for 55 fiber volume composites thermally cycled to 250°C show no significant change from the modulus values of as-fabricated unheated composites. However, composites, thermally cycled to 350°C show a loss in flexural modulus proportional to the number of thermal cycles. Similar behavior was also observed in thermally cycled $FP\text{-}Al_2O_3/EZ33$ Mg composites. After 3000 cycles to 350°C, 55 fiber volume composites degraded to near 90 percent of the modulus values of unheated composite specimens. Also shown in Figure 2 is the modulus values for similar composite specimens isothermally treated at 350°C for 1000 hours, a time equivalent to the cumulative time at temperature for the 3000-cycle test. These specimens did not show any loss inflexural modulus even after 100 hours of exposure. In the cyclic tests, while the specimen length and weight remained nearly the same, the width and thickness of the specimen increased continuously with cycling. In isothermally heated composites, however, no dimensional changes were measured even after 100 hours of exposure of 350°C. The losses in the dynamic modulus of the cycled composites were also found to be proportional to the width and thickness changes.

The room temperature axial tensile strengths of the 55 volume percent composites cycled to 250° or 350°C to a maximum of 3000 cycles are shown in Figure 3. Again, the data points indicate the range and the average

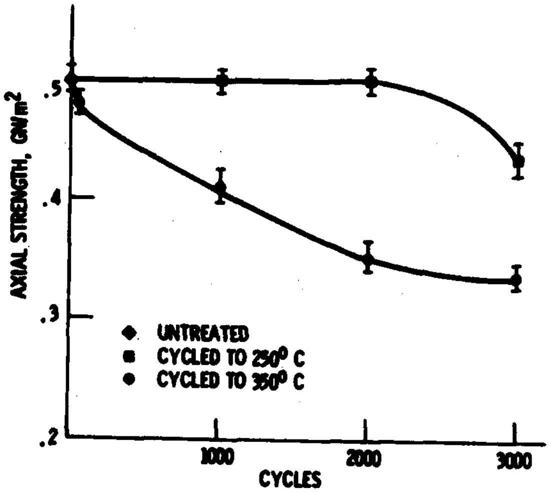


Figure 3 Room temperature axial tensile strengths of 55 volume percent $FP\text{-}Al_2O_3/QH21A$ Mg composites after cycling to 250°C or 350°C for indicated number of cycles. Strengths of untreated composites are shown at zero cycles.

value for typically three tests. The axial strength data for the 55 fiber volume percent composites without thermal treatment are also shown in Figure 3 for a baseline comparison. The strength data, as seen in Figure 3 for the 55 fiber volume percent composite cycled to 250°C shown no appreciable degradation from the unheated composite strength of 0.50GN/m^2 after 2000 cycles. Additional cycling to 3000 cycles resulted in a loss of strength to 0.435 GN/m^2. Similar composites cycled to 350°C showed a gradual loss of strength from the baseline value of 0.50 GN/m^2 after 3000 cycles.

The baseline axial strength data and strength data measured after 1,000 2000, and 3000 cycles from Figure 3 are replotted in Figure 4 against maximum cycle temperature to better illustrate the temperature dependence of the strength degradation. The average strength data of Bhatt, et al.(1) for 55 fiber volume percent FP-Al_2O_3/EZ33 Mg composites in as-received and thermally cycled to 250° or 350°C for up to 3000 cycles are also shown in the figure for comparison. Clearly from Figure 4, no significant strength loss occurred for the 55 fiber volume percent composites when cycled to 250°C, however, a rapid strength loss occurs. This loss appears

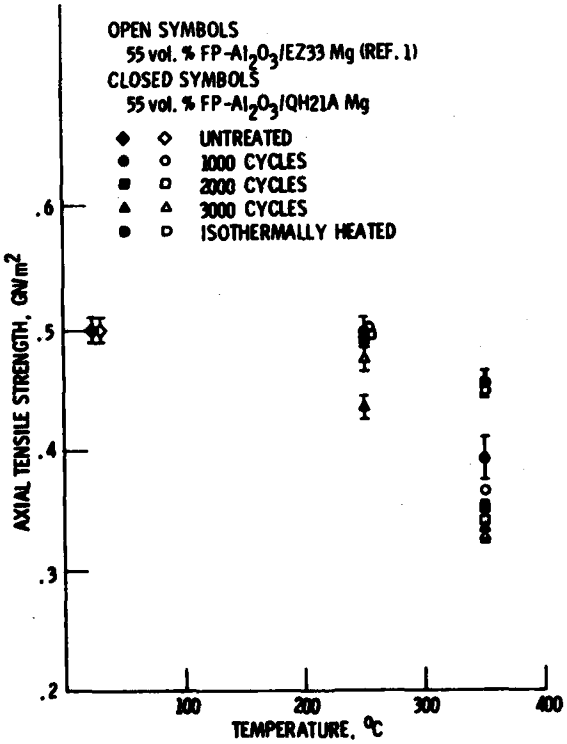


Figure 4 Room temperature axial tensile strengths of FP-Al_2O_3/Magnesium composites cycled 1000, 2000, and 3000 times to indicated temperatures. Strengths of similar composites isothermally heated at 350°C for 100 hours are also shown.

to be more a function of cycle temperature than number of cycles. Also shown in Figure 4 is the range and average value of all the strength data of similar composite specimens which have been isothermally heated at 350°C for 100 hours a time equivalent to the cumulative time-at-temperature for specimens cycled 3000 times. The 55 fiber volume percent composite specimens isothermally heated at 350°C show a 10 percent loss of strength to a value of 0.45 GN/m^2. Strength degradation obtained after isothermally or cyclically heated composites of FP-Al_2O_3/EZ33 Mg or FP-Al_2O_3/QH21A Mg was similar. In all cases, however, the strength values of isothermally heated specimens were equal to or higher than the strength values for thermally cycled composite specimens.

The greater degradation observed for cyclically heated composites is seen as evidence of a mechanism involving more than a simple thermally activated process. If a single thermally activated process were operating, one would expect similar behavior for specimens which were cyclically heated or isothermally heated for an equivalent time at temperature. A likely candidate for this additional mechanism is one which involves the generation of large matrix stresses due to the differences between the thermal expansions of the fiber and matrix during heating. The maximum thermally induced matrix stress, σ_m, in fiber composites can be obtained from the equation derived by Piggott(2).

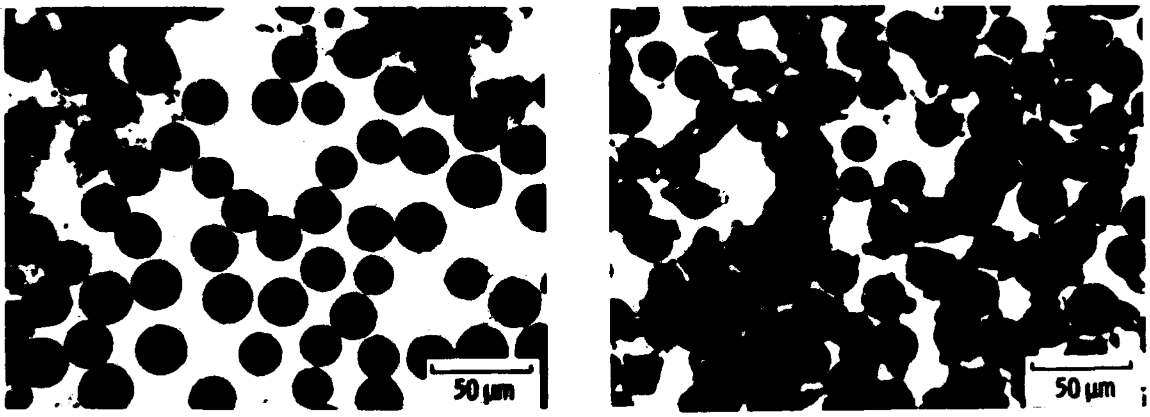
$$\sigma_m = \frac{3\,(\alpha_m - \alpha_f)\,(\Delta T)\,V_f E_m E_f}{E_c + E_f\,(1 + \gamma_m V_m) - \gamma_f E_m}$$

where α is the thermal expansion coefficient, V is the fiber volume fraction, E is the elastic modulus, γ is Poisson's ratio, and ΔT is the cyclic temperature range. The subscripts m, f, and c refer to the matrix, fiber, and composite, respectively. For calculating σ_m, the measured value of E_c = 2.1 GN/m^2 was used along with ΔT = 200° or 300°C, α_m = 25.4 μ/cm/cm/°C, α_f = 5.7 μ/cm/cm/°C, E_f = 379 GN/m^2 and E_m = 44.8 GN/m^2 (3), and γ_f = 0.2, γ_m = 0.33(3,4). The maximum matrix thermal stresses, σ_m, calculated for 55 fiber volume percent composites are shown in Table 1.

Table I

$(T_2 - T_1) = \Delta T$, °C	σ_m, GN/m^2 55 Vol Percent
200	0.174
300	0.262

These stresses are similar to the room temperature yield stress value of 0.17 GN/m^2 for the QH21A magnesium alloy.(3) In the thermally cycled composite therefore plastic deformation will occur during each cycle. The cumulative effect of such repeated deformation has been observed to produce voids in the matrix in highly constrained regions where the matrix deformation cannot be reversed.(5,6) Evidence of similar void growth and matrix cracking in this composite is seen in Figure 5 which shows microphotographs of the 55 volume percent FP-Al_2O_3/QH21A Mg composite before thermal exposure and after cyclic heating to 350°C. The voids here appear to extend to the fiber/matrix interfaces. The presence of voids or cracks at the interface will result in a loss of fiber/matrix bonding.



(a) UNTREATED (b) 3000 CYCLES TO 350°C

Figure 5 Photomicrograph of untreated and thermally cycled composite show effect of cycling to 350°C for indicated number of cycles.

Whether the loss of axial strength results from fiber/matrix debonding or from a weakening of the fiber or the matrix cannot be determined from these micrographs. However, some insight may be obtained from the results of transverse strength of the composite. Figure 6 shows the transverse strength data for 55 fiber volume percent composites thermally cycled to 350°C. It is obvious from this figure that there is no appreciable strength degradation from the untreated composite strength of 0.185 GN/m^2 even after as many as 3000 cycles. Also shown in Figure 6 are the transverse strength data for similar composites isothermally heated at 350°C for 100 hours. As with the data for cycled composites, no appreciable loss in strength was observed at 350°C. This indicates that matrix strength loss may not be a major factor for the strength degradation of thermally cycled composites.

To obtain further understanding about the cause of strength degradation, fracture surfaces of thermally cycled composites were examined. Typical fracture surfaces of the untreated and thermally cycled composites for 1000 and 3000 cycles are shown in Figure 7. This figure indicates fiber debonding and pull-out in thermally cycled composites. Since isothermally heated composites showed no voids or fiber pull-out, we associate the gradual strength decrease with the number of cycles as shown in Figure 3 with debonding of the fiber and the matrix due to growth voids at the fiber matrix interface.

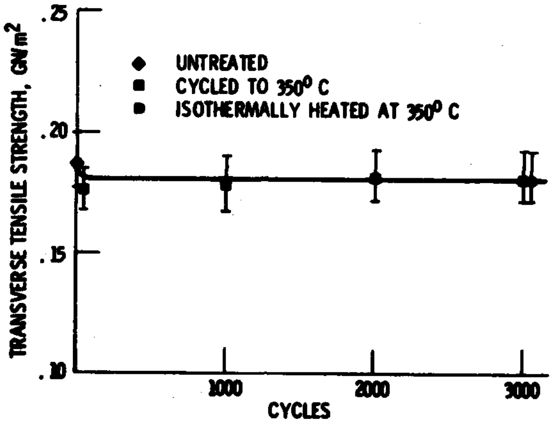
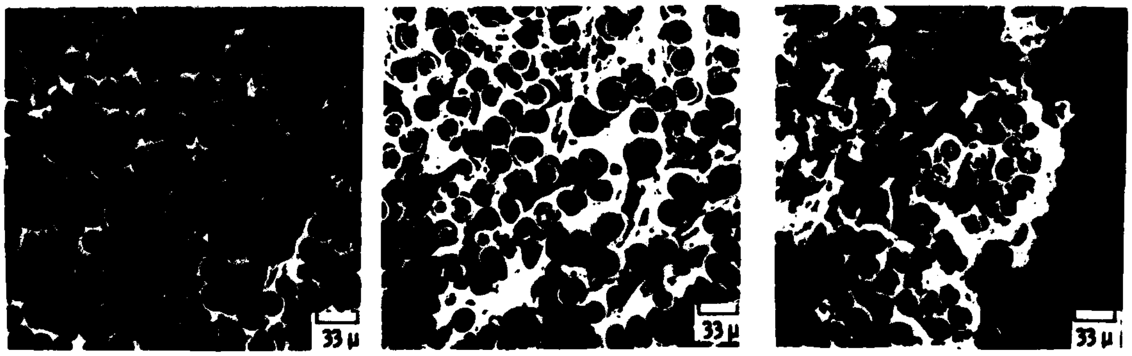


Figure 6 Room temperature transverse tensile strengths of FP-Al_2O_3/ QH21A Mg composites after cycling to 350°C for indicated number of cycles. Strengths of similar composites isothermally heated at 350°C for 100 hours are shown. Strengths of untreated composites are shown at zero cycles.



UNTREATED 1000 CYCLES TO 350° C 2000 CYCLES TO 350° C

Figure 7 Fracture surface of untreated and thermally cycled FP-Al_2O_3/ QH21A Mg composites stressed in axial direction showing effect of cycling to 350°C for indicated number of cycles.

Summary

The effects of cyclic and isothermal heat treatments on the axial and transverse tensile strength and axial moduli of 55 volume percent FP-Al_2O_3/QH21A Mg composites have been evaluated to understand the cause of thermally induced strength degradation and to help determine the limiting use conditions for these composites. Specific findings are as follows:

1. Thermal cycling of FP-Al_2O_3/QH21A Mg composites to 250°C for 2000 cycles did not cause any appreciable room temperature strength or modulus loss compared with baseline data for untreated composites. In contrast, composites thermally cycled to 350°C showed considerable loss in both axial strength and modulus. No appreciable loss in room temperature transverse tensile strength of composites was measured after thermal cycling.

2. Measurement of the transverse strength and fractographic analysis of thermally cycled composites indicated interface void formation and matrix cracking and fiber debonding as prime contributors to observed strength and modulus losses. These results are consistent with degradation mechanism based on thermal-induced stresses in the matrix.

3. No appreciable loss in the axial dynamic modulus of these composites was observed after isothermal exposure. Small strength losses observed in isothermally heated composites were attributed to matrix softening.

4. The high temperature mechanical properties were not measured for the composites in this study. This would be required to properly design using FP-Al_2O_3/QH21A composites. However, aside from the usual matrix softening at higher temperatures, this study indicates that we would not expect additional axial strength or modulus degradation resulting from either isothermal or cyclic exposure below 2000 cycles or 250°C.

References

(1) Bhatt, R. T. and Grimes, H. H., "Thermal Degradation of the Tensile Properties of Unidirectionally Reinforced FP-Al_2O_3/EZ33 Mg Composities," NASA TM-82817, 1982.

(2) Piggott, M. R., "Load Bearing Fiber Composites," pp. 208, Pergamon Press, Oxford, NY, 1980.

(3) Metals Handbook, Vol. I, "Properties and Selection of Metals," American Society for Metals, Metals Park, OH, 1961.

(4) Bacon, J. F., Prewo, K. M., and Veltri, R. D., "Glass Matrix Composities-II-Alumina Reinforced Glass," pp. 753-769, in Proceedings of the 1978 International Conference on Composite Materials, Noton, B., Signorelli, R., Street, K., and Phillips, L., Editors, AIME, New York, NY, 1978.

(5) Shahinian, P., "Thermal Fatigue of Aluminum-Boron Composites," SAMPE Quart., Vol. 2, 1970, pp. 28-35.

(6) Kim, W. H., Koczak, M. J., and Lawley, A., "Effects of Isothermal and Cyclic Exposures on Interface Structure and Mechanical Properties of FP -Al_2O_3/Aluminum Composites," New Developments and Applications in Composites, Doris Kuhlmann-Wildorf and W. C. Harrigan, Editors, Proceedings of a Symposium sponsored by the TMS-AIME Physical Metallurgy and Composites, 1979, pp. 40-53.

THE AERODYNAMIC INFLUENCES OF ROTOR BLADE TAPER, TWIST, AIRFOILS AND SOLIDITY ON HOVER AND FORWARD FLIGHT PERFORMANCE

GENE J. BINGHAM, MR.
STRUCTURES LABORATORY
US ARMY RESEARCH & TECHNOLOGY LABORATORIES (AVRADCOM)
LANGLEY RESEARCH CENTER, HAMPTON, VA 23665

Introduction

The use of composites for helicopter rotors and the development of new airfoils has provided new opportunities in blade design. With composites, the fabrication of non-rectangular blades with variations in airfoil sections along the blade radius is practical. The advantages of non-rectangular rotor blades and variations in twist distribution and airfoils were considered in the late 1940's (refs. 1 to 5). These studies addressed rotors operating at relatively low tip speeds and forward flight velocities. For the configurations considered, increasing the twist from zero to -12 degrees provided thrust increases up to 4 percent and tapering the blades from root to tip increased the thrust an additional one or two percent in hovering flight. The analysis of the influences of twist and planform was not extended to forward flight.

Because current rotor tip speeds and flight speeds are significantly higher than those of references 1 to 5 and that rotor analyses techniques have been improved, a study has been performed to consider the influences of blade taper, point of taper initiation, twist, solidity, and airfoil on rotor efficiency in hover and forward flight. The study began with the design of an advanced rotor for the UH-1 helicopter. The initial design goal was to reduce hover power required by 8 percent without degrading forward flight performance. This reduction was to be accomplished with an aircraft gross weight of 8050 pounds while operating at an altitude of 4000 feet and a temperature of 95°F.

The study indicated that the design goal could be exceeded. Based on this result, models of the baseline and advanced blade have been evaluated in the Langley V/STOL wind tunnel (refs. 6 and 7) and the analytical study has been extended to other helicopter configurations within the US Army inventory. This paper is to describe the design philosophy applied.

The influence of blade planform and twist on rotor performance are considered first for hover and then for forward flight. These influences initially are made independent of airfoil characteristics; after the influences of blade geometry are described, the airfoil requirements are addressed.

Discussion

Analytical Approach

The analytical results presented in this paper have been developed by using a simple hover and forward flight analyses. The programs were selected for high computer productivity and do provide qualitative results. The hover analysis combines the momentum theory and the blade-element theory (ref. 5). The forward flight analysis is tailored after references 8, 9, and 10; the rotor characteristics considered include thrust, profile-drag power, total power, flapping, rolling and pitching moments, direction of the resultant force vector, and the harmonic contributions of each rigid blade on the rotor to the shear-force input to the hub. Both the hover and forward flight analyses apply two-dimensional airfoil data tables in performing the blade element integrations. The Rotorcraft Flight Simulation Computer Program, C81 (ref. 11), was applied in the final forward flight analysis.

The initial steps in the analysis are to evaluate the influences of blade twist and taper on rotor performance and to identify the desired airfoil aerodynamic characteristics. The influence of airfoil stall and drag divergence are avoided in these steps by applying an "ideal" airfoil. The aerodynamic characteristics of this airfoil include a constant lift curve slope for all angles of attack and a constant drag coefficient (selected to be 0.0070) at all angles of attack and Mach numbers. The blade section lift coefficients of the blade-element computations will be equal to those of a "conventional" airfoil operating in the linear lift coefficient-angle of attack range.

Hover Analysis

The influences of blade taper ratio and the radial station at which taper begins is presented in figure 1. The ordinate is figure of merit $\left(F.M. = 0.707\ C_T^{3/2}/C_Q\right)$ and the abscissa indicates the radial station at which blade planform taper begins; the blades would be rectangular to the station where taper begins and the completely rectangular blade would be represented at a value of 1.0. As previously noted, a UH-1 type rotor blade is considered; and for this figure the area solidity (σ = .0464) is constant.

The most significant change in figure of merit observed results from the change in radial station where taper is initiated. The increase is as great as 12 percent (F.M. = 0.80 compared to 0.715) by changing from a rectangular blade to a blade tapered from 0.2R. For each taper ratio, nearly one-half of the increases shown results from a blade which is rectangular to 0.90R and then tapered to the blade tip; moving the taper initiation inboard of 0.50R has little influence on figure of merit at some taper ratios. The influence of taper ratio tends to increase as the onset of taper is moved inboard. For example, at 0.50R increasing taper ratio from 3 to 9 results in a figure of merit increase of about 2.6 percent (F.M. - 0.775 to 0.795). The indicated favorable influences of location of taper initiation and taper ratio result from (1) the increased inboard loading of the rotor which decreases induced torque and (2) the reduction in chord in the outboard regions which decreases profile torque. The increased inboard loading decreases induced torque because the induced torque is a function of the product of airfoil section lift coefficient, chord, and radius squared; therefore, an advantage exists in distributing the greater lift coefficient and chord at the inboard radial stations. The reduction in chord in the outboard regions decreases the profile torque because the profile torque is a function of the product of airfoil section drag coefficient, chord, and radius squared; therefore, an advantage can be seen in reducing blade chord at the outboard radial stations.

As will be discussed later, maintaining a constant area solidity (as for fig. 1) may not be feasible in the final process of combining planform,

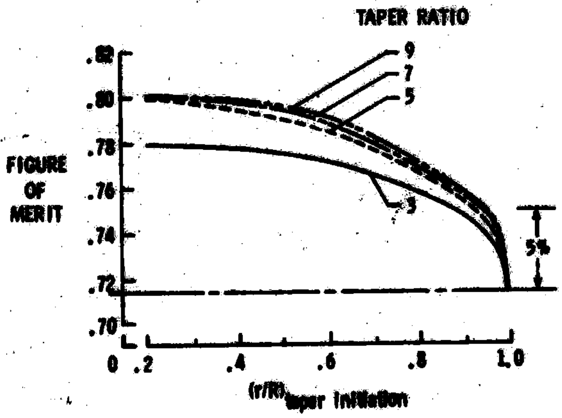


Figure 1.-Influence of location of initiation of taper and taper ratio on figure of merit, σ = 0.0464, θ = -10.6°, 8030 lbs, 4000'/95°.

twist, and airfoils due to airfoil section stall. Therefore, figure 2 is defined to consider the influence of a constant thrust weighted solidity. In providing improvements for existing rotor configurations, maintaining the thrust weighted solidity may not be necessary but should be the maximum solidity required. The trends are similar to those with constant area solidity but the differences in figure of merit due to position of taper initiation and taper ratio are reduced in magnitude. In this case, the increases in figure of merit are as great as 6 percent which is only one-half that indicated with a constant area solidity. An analysis of these results (where thrust is considered to be constant) reveals that for a given position of taper initiation and taper ratio the induced torque is unchanged by variations in solidity. The induced torque is unchanged since the radial distribution of lift for a given configuration is unchanged; at a given radial station, an increase in chord is inversely proportional to the decrease in lift coefficient, so the product of lift coefficient and local chord is unchanged. Therefore, the differences in figure of merit with changes in solidity are due entirely to differences in profile torque which result from differences in local airfoil chord. As will be shown in the discussion of the airfoil requirements section of this paper, the solidity applied depends on the combined influences of planform, twist, and the aerodynamic characteristics of the airfoil section selected.

The influence of changes in linear twist are presented in figure 3 for a taper ratio of 3. An increase in twist would, of course, result in a general increase in inboard loading and corresponding decrease in outboard loading to have an influence similar to that described for the previous figures; for a given planform, only the induced torque would be influenced for the "ideal" airfoil which has a constant drag coefficient. The most significant increase in figure of merit due to twist is indicated for the

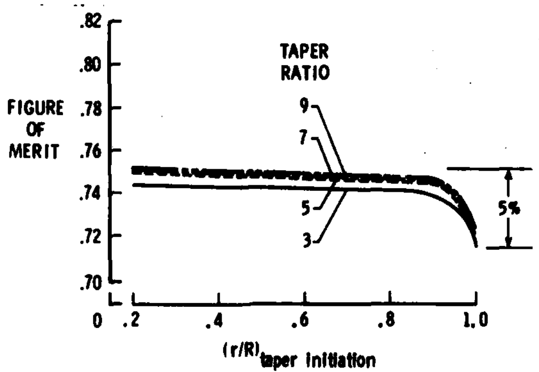


Figure 2.-Influence of location of initiation of taper and taper ratio on figure of merit, $\sigma_T = 0.0464$, $\theta = -10.6°$, 8050 lbs, 4000'/95°.

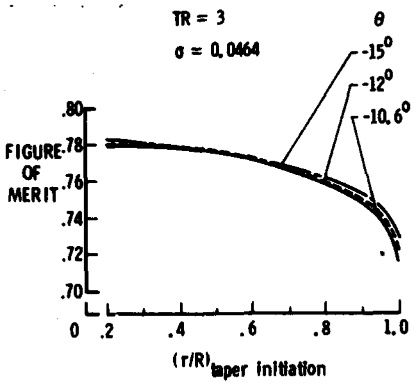


Figure 3.-Influence of linear twist on figure of merit, σ = 0.0464, 8050 lbs, 4000'/95°.

rectangular planform (r/R = 1.0); the increase in twist from -10.6° to -15° increases the figure of merit about 2 percent. As the taper initiation position is moved inboard, the influence of increased twist is reduced.

Forward Flight Analysis

The influence of blade taper and the radial station at which taper is initiated on forward flight performance is indicated in figure 4. The ordinate is torque coefficient and the abscissa again defines the radial position where taper is initiated. The forward flight velocity considered is 110 knots (near maximum for the UH-1 in level flight); the area solidity is constant, inflow is uniform, and the other parameters are the same as those for figure 1. Similar to the case for hover, a significant change in torque coefficient results from changes in radial station where taper is initiated; the decrease is as great as 16 percent by initiating taper at 0.2R. Unlike the previously discussed case for hover (fig. 1), only a small part of the total decrease results from changing the blade from fully rectangular to a configuration with taper initiated at 0.9R and the advantage of moving the taper initiation location inboard of 0.5R is greater in forward flight than in hover. As for the hover case, increases in taper ratio from 3 to 9 provide several percent decreases in torque coefficient. The favorable influences are due to the inboard loading and the reduction in chord in the outboard regions.

The results of the forward flight analysis with constant thrust weighted solidity are presented in figure 5. As for hover, the increased

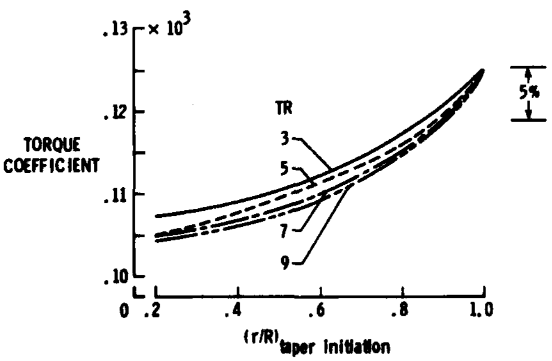


Figure 4.-Influence of location of initiation of taper and taper ratio on torque coefficient, $\sigma = 0.0464$, $\theta = -10.6°$, 8050 lbs, 4000'/95°, 110 knots.

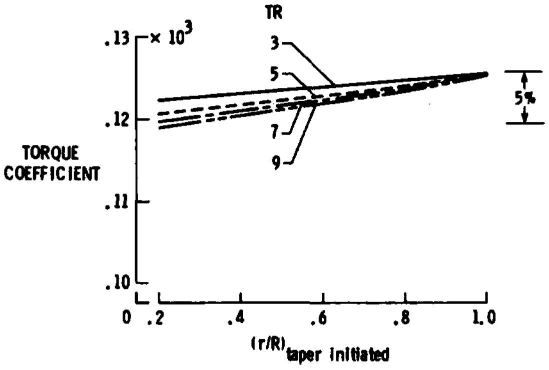


Figure 5.-Influence of location of initiation of taper and taper ratio on torque coefficient $\sigma_T = 0.0464$, $\theta = -10.6°$, 8050 lbs, 4000'/95°, 110 knots.

blade chord (for a given taper initiation location and a given taper ratio) results in an increase in torque coefficient (compare figs. 4 and 5). The maximum torque advantage of planform change observed with constant area solidity (16 percent) is reduced to only 5 percent with constant thrust weighted solidity. Therefore, it is appropriate to consider the solidity requirements. It is well known that the solidity should be sufficient to (1) satisfy the hover design lift requirement, (2) avoid retreating blade stall at the forward flight design condition, and (3) avoid stall in maneuvers. Considering that the airfoil section lift coefficients required in hover are significantly less than those required in forward flight, the steady forward flight and maneuvering lift coefficient requirements become the controlling factors in the design.

Prior to considering the lift coefficient requirements and thus the airfoil requirement, it is desirable to consider the influence of changes in linear twist in forward flight. The twist influence at a taper ratio of 3 is indicated in figure 6 for constant area solidity. Unlike the results for hover (figure 3) the increase in twist for the rectangular blade (r/R = 1.0) had little influence on the configuration considered. The greatest influence of twist occurred with taper initiation near 0.5R; the reasons for these results are not readily apparent.

Airfoil Requirements

The hover and forward flight analysis discussed previously have shown, by example, the potential benefits of changes in blade planform and twist.

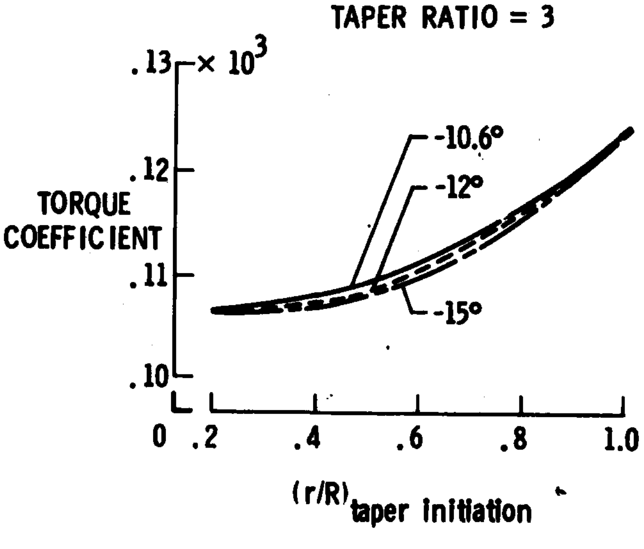


Figure 6.-Influence of linear twist on torque coefficient, $\sigma = .0464$, 8050 lbs, 4000'/95°, 110 knots.

The benefits observed at a constant area solidity were decreased when a constant thrust weighted solidity was maintained. Therefore, it is appropriate to now consider the airfoil requirements in light of the blade configurations previously discussed. This can be accomplished by applying the forward flight analytical program used in the earlier discussion; the forward flight analysis defines the upper and lower limits of lift coefficient requirements. As previously noted, the influences of airfoil stall and drag divergence can be avoided by applying the "ideal" airfoil so the first order lift coefficient-Mach number requirements for different configurations can be identified. Therefore, the airfoil lift requirements can be defined. The lift coefficient requirements will be unchanged for actual airfoil configurations if the rotor blade sections are operated in the linear range at all Mach numbers; only the drag coefficients will be different with different airfoils. For maneuvering flight the analysis becomes more complex and is not considered in this paper. In reviewing the airfoil requirements, the same 8050 pounds gross weight aircraft at 4000'/95° is considered.

The influence of position of initiation of taper on lift coefficient requirements is shown in figure 7. Section lift coefficients are presented

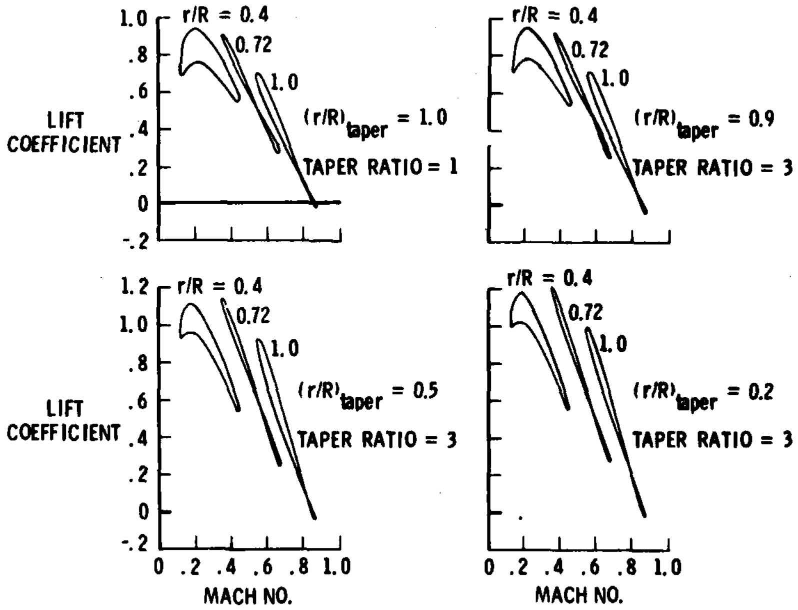


Figure 7.-Influence of position of initiation of taper on lift coefficient, σ = 0.0464, θ = -10.6°, 8050 lbs, 4000'/95°, 110 knots.

for three radial stations (0.4R, 0.72R, and 1.0R) as a function of section Mach number. Plots are provided for a rectangular rotor and for rotors with a taper ratio of 3 and with taper initiated at 0.9R, 0.5R, and 0.2R; area solidity is constant.

The inboard movement in radius at which taper is initiated results in a general increase in lift coefficient required at all radial stations. The primary increase is on the retreating side because the dynamic pressure is lower than on the advancing side. Changing from the rectangular blade to the blade tapered from 0.2R results in an increase in the lift coefficient required by as much as 0.3.

The influence of increasing the taper ratio from 3 to 5 or 9 are shown on figure 8; again, the area solidity is constant. With taper initiated at 0.9R, the lift coefficients are only slightly different from those for the rectangular configuration. However, with taper initiated at 0.5R and 0.2R,

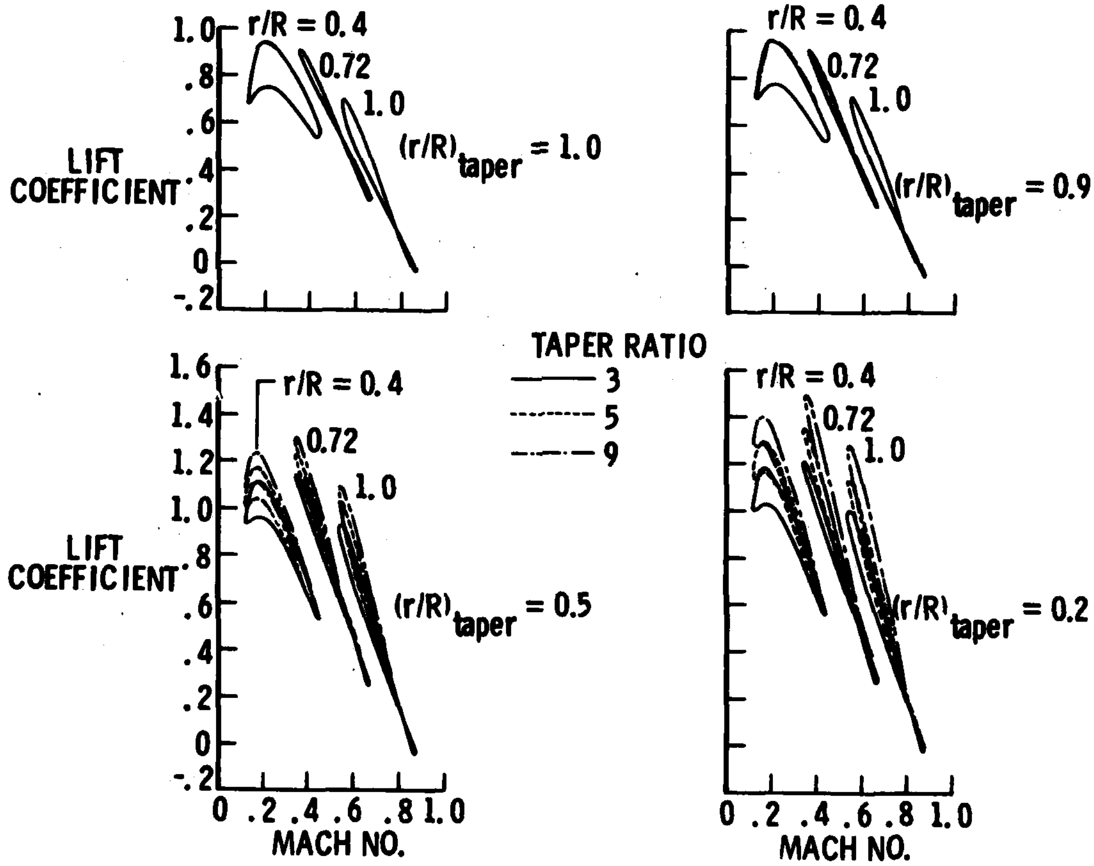


Figure 8.-Influence of taper ratio on lift coefficient, $\sigma = 0.0464$, $\theta = -10.6°$, 8050 lbs, 4000'/95°, 110 knots.

the increases in required lift coefficient are as large as those resulting from the changes in taper initiation observed in figure 7. It is of interest to note here, at 1.0R, satisfying both a drag divergent Mach number of 0.87 and a maximum lift coefficient near 1.3 for taper initiation at 0.2R is beyond the capability of known airfoil sections; also, it should be noted that the position at which taper can be initiated and the blade taper ratio can be controlled by the maximum lift coefficients of the available airfoils.

As shown in figure 9, a reduction in the maximum lift requirement at the blade tip (r/R = 1.0) can be accomplished by increasing the linear twist of the blade. In the example shown, the increase is from -10.6 degrees to -15 degrees. The tip section lift coefficients are decreased at all Mach numbers. At 0.72R the twist has little influence on the operational lift coefficient. At 0.4R, the increased twist increases the inboard lift coefficients to result in an increase in inboard loading. The increase in twist results in requirements more consistent with current airfoil capability than those observed in figure 8 and permits adjustments in requirements to match the maximum lift coefficients of available airfoils.

The influences of blade solidity on the lift coefficient are shown in

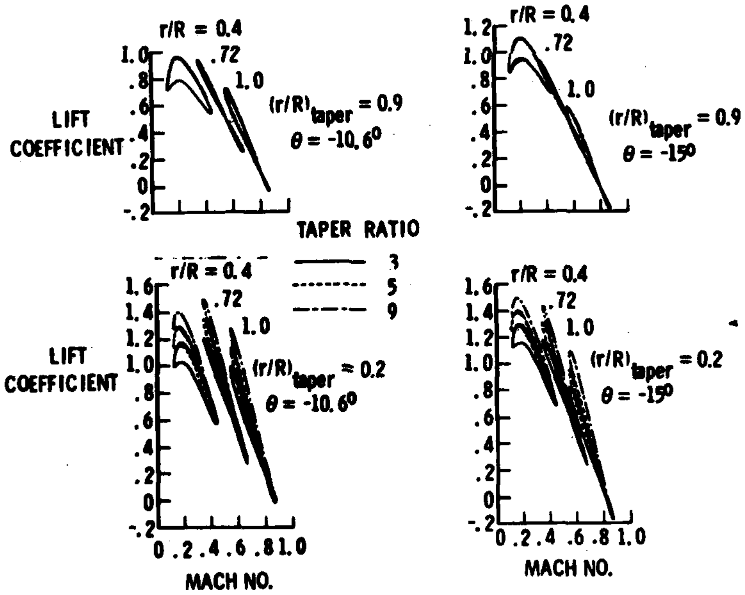


Figure 9.-Influence of taper ratio and linear twist on lift coefficient, σ = 0.0464, θ = -10.6°, 8050 lbs, 4000'/95°, 110 knots.

figure 10 where curves are presented at constant area and thrust weighted solidity. The plot at the upper left is once again for a rectangular blade and the other three are for a taper ratio of 3 with taper initiation locations at 0.9R, 0.5R, and 0.2R. The solidity increase results in reductions in lift coefficient. The greatest difference is indicated for the case with the most inboard taper initiation. It should be recalled, however, that the increase in solidity resulted in a reduction in the benefits of inboard taper initiation and taper ratio. The point to be recognized now is that the ability to obtain the improvements in the hover figure of merit (figs. 1 and 2) and the forward flight torque reductions (figs. 4 and 5) discussed earlier will depend on relationships between the lift coefficient-Mach number requirements and the lift coefficient-Mach number capability of the selected airfoil(s). The full advantages (for hover and forward flight) of the more inboard taper and the greater taper ratios can be realized only if the airfoils will satisfy the higher lift coefficient requirements at all radial stations and if drag divergence Mach number at all lift coefficients is not exceeded. As the characteristics of the airfoil are reviewed relative to the requirements, the selection of the initiation of taper, taper ratio, twist and solidity can be accomplished for approximating the best possible design. Also, the final solidity selection process should include consideration of maneuvering performance.

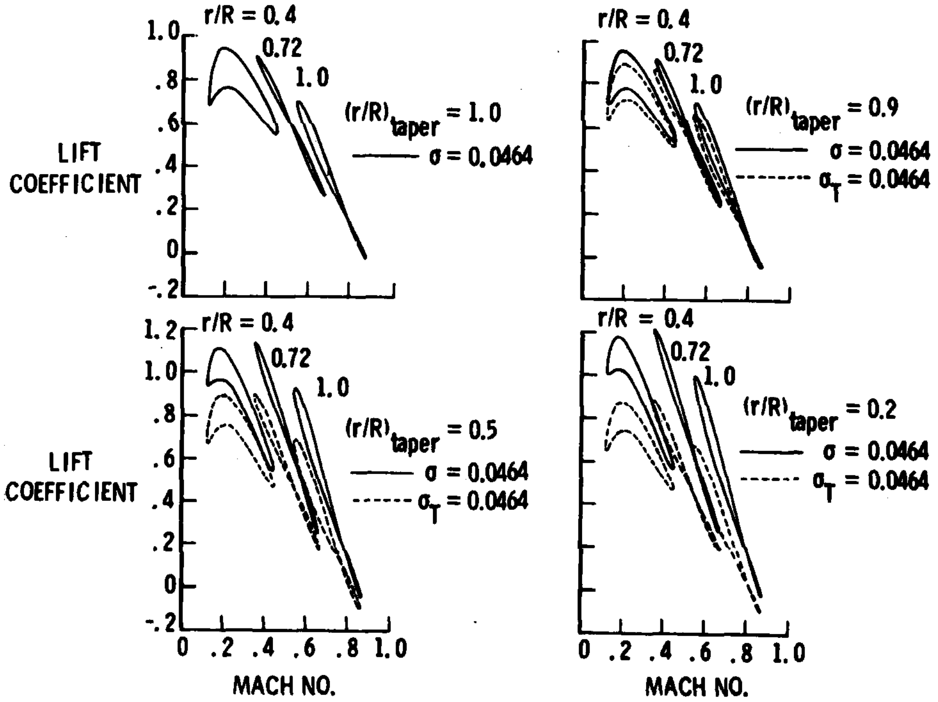


Figure 10.-Influence of solidity on lift coefficient, taper ratio = 3.0, θ = -10.6°, 8050 lbs, 4000'/95°, 110 knots.

Design Application

As previously noted, the design philosophy described herein has been applied to a UH-1 type rotor. For this design the goal was to provide an 8 percent reduction in hover power required when operating with 8050 pounds gross weight at 4000 feet altitude and 95 degrees Fahrenheit; the blade radius of 24 feet was to be maintained and forward flight performance was not to be degraded. The resulting configuration is shown on figure 11. along with the standard rectangular blade. The new configuration has a linear twist of -14 degrees, is rectangular to 50 percent radius, and then is tapered to the blade tip with a taper ratio of three. The area solidity is approximately 5 percent greater than that of the standard blade and the thrust and torque weighted solidity are reduced approximately 18 and 25 percent, respectively. Three airfoils (ref. 12) are distributed along the radius with a 12 percent thick section to 0.8R, a 10 percent thick section is located in the thickness transition region between 0.8R and 0.9R and an 8 percent thick section is applied from 0.9R to the blade tip.

The airfoils, planform, and twist selection began by considering the requirements suggested by figure 12 which displays the lift coefficient-Mach number requirements at the gross weight, altitude, and temperature addressed in the design. The forward flight velocity (110 knots) is near the maximum for these conditions in level flight. The maximum lift coefficient and the drag divergence Mach number of the NACA 0012 airfoil have been added to figure 12 for comparison. This figure suggests that sufficient lift coefficient is available so that the inboard regions might be more heavily loaded and that the drag divergence Mach number for the advancing region needs to be increased.

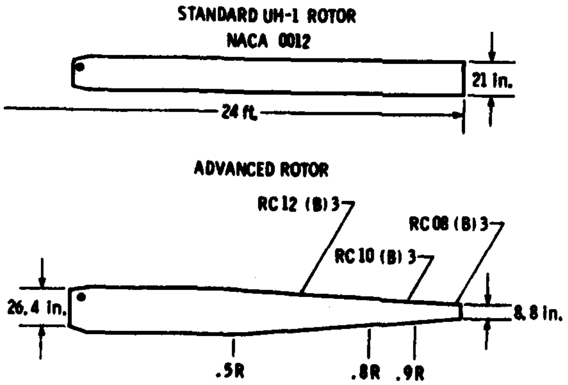


Figure 11.-Planform of standard and advanced rotor.

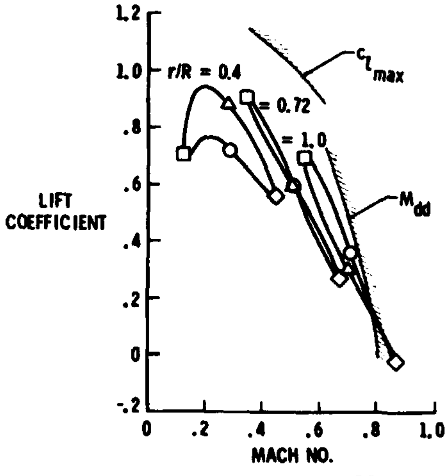


Figure 12.-Operational lift coefficients for UH-1D rotor, 8050 lbs, 4000'/95°, 110 knots.

The baseline and advanced rotor configurations have been evaluated in the NASA Langley V/STOL tunnel on a UH-1 quarter-scale model complete with fuselage (refs. 6 and 7). Data from the tests are presented in figure 13 to indicate the correlation between the hover analysis program results (applied in the earlier discussion of this paper) and the hover data from the V/STOL wind tunnel. In the analysis, two-dimensional airfoil data obtained in the Langley 6- x 28-inch transonic wind tunnel was applied. Data for the NACA 0012 airfoil was applied in the standard blade analysis and the airfoils noted on figure 11 were applied as distributed along the blade radius. The design thrust coefficient corresponding to the gross weight and altitude/temperature is 0.0032. The experiment is in agreement with the analysis and provides validation of the design philosophy presented in this paper. At the design thrust coefficient, the advanced rotor operated at a torque coefficient about 10 percent below that of the standard blade.

In forward flight, the percentage improvement was greater than in hover (figure 14). Although not shown on this figure, the experimental results are consistent with analysis. The power reduction due to the advanced blade is about 19 percent at the gross weight of this design. This corresponds to a range increase of about 11 percent.

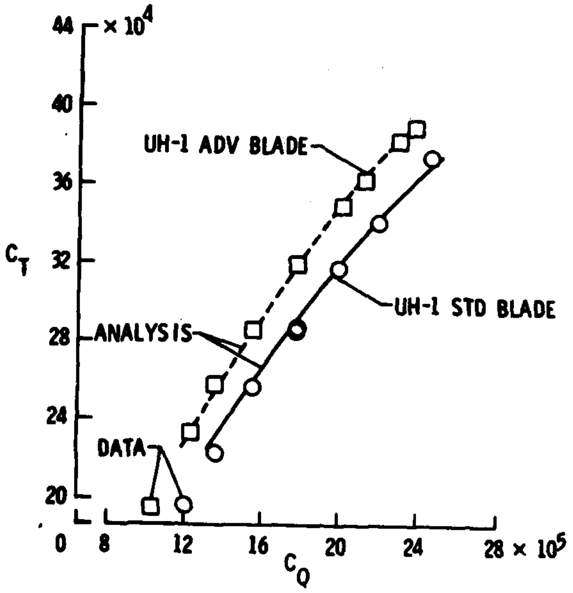


Figure 13.-Comparison of hover wind tunnel data and hover analysis.

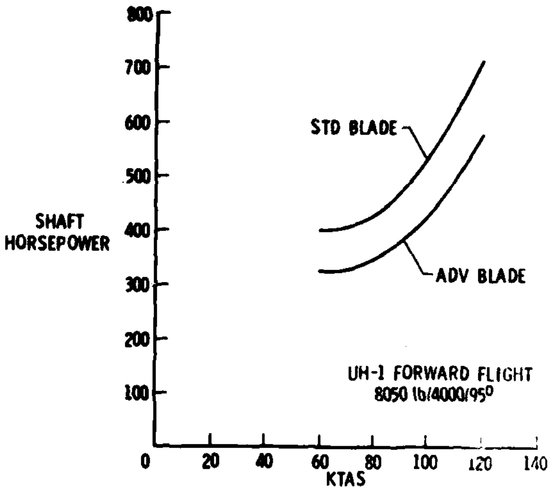


Figure 14.-Comparison of predicted forward flight performance of standard and an advanced UH-1 main rotor blade.

Concluding Remarks

It has been determined that the use of non-rectangular rotor blades can significantly improve the rotor performance, even if it is necessary to maintain a constant thrust weighted solidity. The preliminary design can be accomplished by applying "ideal" airfoil characteristics in simple hover and forward flight analyses. The "ideal" airfoil can also approximately define the airfoil lift coefficient requirements to lead to the final airfoil selection. The final design must then be evaluated with the characteristics of the airfoils to be applied. With the airfoil characteristics applied, adjustments may prove necessary in blade twist, planform and solidity.

Although not discussed in this paper, other performance requirements must be considered in the final design. Included would be autorotation characteristics which would depend in part on the mass distribution of the blade. Maneuver performance would also be addressed which would certainly depend on thrust weighted solidity. It is clear, however, that the development of improved airfoils to meet design requirements provides new opportunities in blade design.

References

1. Gustafson, F. B.: Effect on Helicopter Performance of Modification in Profile-Drag Characteristics of Rotor Blade Airfoil Sections, NACA ACR L4H05, 1944.

2. Gustafson, F. B. and Gessow, Alfred: Effect of Rotor-Tip Speed on Helicopter Hovering Performance and Maximum Forward Speeds. NACA ACR L6A16, 1946.

3. Gessow, Alfred: Effect of Rotor-Blade Twist and Plan-form Taper on Helicopter Hover Performance, NACA TN 1542, 1948.

4. Gessow, Alfred: Flight Investigation of Effects of Rotor-Blade Twist on Helicopter Performance in the High Speed and Vertical-Autorotative Descent Conditions, NACA TN 1666, 1948.

5. Gessow, Alfred and Myers, Garry C., Jr.: Aerodynamics of the Helicopters. Frederick Unger Publishing Company, New York, 1952.

6. Berry, John D.: Quarter Scale Testing of an Advanced Rotor System for the UH-1 Helicopter, Paper No. 81-49, 37th Annual Forum of AHS, 1981.

7. Hoad, Danny R. and Conner, David A.: Acoustic Performance Evaluation of an Advanced UH-1 Helicopter Main Rotor System, Paper No. 81-58, 37th Annual Forum of AHS, 1981.

8. Gessow, Alfred: Equations and Procedures for Numerically Calculating the Aerodynamic Characteristics of Lifting Rotors. NASA TN-3747, October 1956.

9. Gessow, Alfred and Crim, Almer D.: A Method for Studying the Transient Blade-Flapping Behavior of Lifting Rotors at Extreme Operating Conditions, NACA TN-3366, January 1955.

10. Gessow, Alfred and Tapscott, Robert J.: Tables and Charts for Estimating Stall Effects on Lifting-Rotor Characteristics. NASA TND-243, May 1960.

11. Van Gaasbeek, J. R.: Rotorcraft Flight Simulation, Computer Program C81, USARTL-TR-77-54B, 1979.

12. Bingham, Gene J., Noonan, Kevin W., and Jones, Henry E.: Results of an Investigation of Several New Rotorcraft Airfoils as Related to Airfoil Requirements. NASA Conference Publication 2046, 1978.

DIELECTRIC WAVEGUIDE ANTENNA (U)

*JOHN BOROWICK, WILLIAM BAYHA
COMBAT SURVEILLANCE & TARGET ACQUISITION LABORATORY
USA ERADCOM

RICHARD A. STERN, RICHARD W. BABBITT
ELECTRONICS TECHNOLOGY & DEVICES LABORATORY
USA ERADCOM

Introduction

This paper will describe the novel design of a non-metallic dielectric waveguide frequency scan antenna. This antenna is capable of frequency scanning a beam over a spatial angle four times greater than that obtained with a conventional slotted metal waveguide antenna design. This is done with the same percentage change in frequency and there is no metallization required of the radiating aperture. The entire radiating aperture and the transmission line are an integral, homogeneous material; no individual assembly of radiating elements to the transmission line is necessary. The absence of metallization not only means that the fabrication procedure has been considerably simplified, but also that the conductor ohmic losses are eliminated.

The antenna is made by cutting grooves into the surface of a dielectric waveguide. The grooves act as radiating elements. This paper will describe the design and test results of several line source antennas that were investigated at Ka Band, for applications in the millimeter wave frequency range.

Background

Antenna beam scanning can be accomplished by several means. Often the simplest expedient is to mechanically scan the antenna by physically moving the entire antenna structure. For system applications where fast and precise beam steering is required, inertialess beam scan is used. Inertialess beam scan is usually accomplished electronically by altering the phase across the radiating aperture with discrete phase shifting components or by altering the frequency whereby an inherent phase is attained between individual radiating elements. All the above techniques are used at microwave frequencies. For millimeter wave frequencies (i.e., 30-300 GHz) mechanical scan is used in virtually all systems. Inertialess scanning at millimeter waves is difficult because of the impracticality in size of the components that would be used for beam steering. It is readily apparent to the millimeter wave antenna engineer that new concepts are needed for beam scanning at millimeter waves. (1)

The Combat Surveillance and Target Acquisition Laboratory has recognized the need for new antenna techniques and has been working under a joint internal Army effort with the Electronics Technology and Devices Laboratory to develop a dielectric waveguide line source antenna. This antenna is unique in that the transmission line and the antenna aperture are an integral, homogeneous structure. The dielectric waveguide transmission line is simply grooved with a saw blade so as to form the radiating elements. This is an inherently simple, low cost fabrication process. The integral design of this transmission line - radiating aperture implies a potential for conformal mounting to a curved surface. When these highly desirable fabrication attributes are noted, along with the significant fact that there is a factor of four increase in the spatial scan sensitivity, this design becomes very interesting.

Dielectric Waveguide

The cross section of the dielectric waveguide was designed using the analysis of Marcatili. (2) The cross section dimensions used were a = 0.052" and b = 0.070", which allows only propagation of the lowest order E_{11}^{y} mode, which is shown in Figure 1. The dielectric waveguide was tapered on both ends so as to permit a transition to standard Ka Band waveguide for test purposes. In particular, each dielectric waveguide had an H plane taper at both ends, which served as the transition between the standard TE_{10} metal waveguide mode to the E_{11}^{y} dielectric waveguide mode. By choosing magnesium titanate, a high dielectric-constant material for which the index of refraction $n = 4$, nearly all of the power flow is confined to the waveguide. The transverse wavenumbers, k_x and k_y, are obtained in our case from the transcendental equations

$$k_x a = \pi - 2 \tan^{-1}\left[(n^2 - 1)k_o^2/k_x^2 - 1\right]^{-\frac{1}{2}}, \qquad (1)$$

and

$$k_y b = \pi - 2 \tan^{-1} \left[n^4 (n^2 - 1) k_o^2 / k_y^2 - 1 \right]^{-\frac{1}{2}}, \qquad (2)$$

respectively, where k_o is the free-space wavenumber corresponding to the operating frequency. The propagation wavenumber k_z is obtained from

$$k_z = \left[n^2 k_o^2 - k_x^2 - k_y^2 \right]^{\frac{1}{2}} . \qquad (3)$$

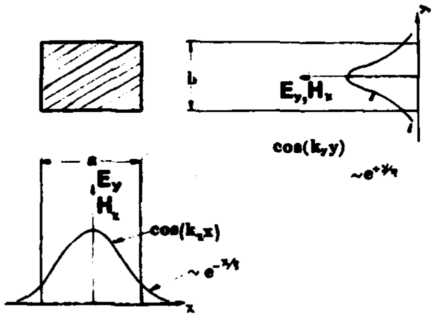


FIGURE 1. E^y_{11} DIELECTRIC WAVEGUIDE MODE

Calculated values of the guide wavelength, $\lambda_g = 2\pi/k_z$, are plotted in Figure 2. A knowledge of λ_g is required in order to proceed to the next step of the design procedure, i.e., a determination of radiating slot spacing, which will, in turn, determine the beam pointing direction.

Dielectric Waveguide Antenna

It is evident from Figure 1 that the E_y and H_x field components are not zero at the edge of the waveguide. Thus there exists a surface wave at the dielectric-air interface. If perturbations are spaced at some periodic interval d along the length of the guide, a radiation field will be created. In our case the perturbations consist of transverse slots cut into the dielectric, of width w and depth t. As indicated in Figure 3, the E-plane beam peak appears at an angle θ with respect to the normal, given by the expression

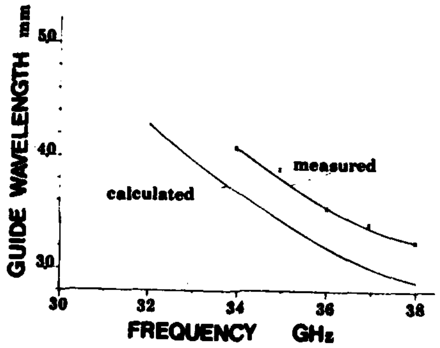


FIGURE 2. GUIDE WAVELENGTH VS. FREQUENCY

$$\sin \theta = \lambda_o/\lambda_g + m\, \lambda_o/d \ , \tag{4}$$

where λ_o is the free-space wavelength corresponding to the operating frequency. It is evident that the beam may be scanned in space by changing the frequency. The quantity m in Eq. (4) is limited to $m = -1, -2, -3, \ldots$ corresponding to spatial harmonics in the visible space $(-\pi/2 \leq \theta \leq \pi/2)$. For maximum coupling (3), as well as single beam operation from backfire to endfire, the value $m = -1$ is chosen. The value of d is then limited by the condition

$$\left| \lambda_o/\lambda_g - \lambda_o/d \right| \leq 1. \tag{5}$$

The slots were cut into a slab of dielectric by a precision-controlled diamond dicing saw. The slab was then machined to the finished waveguide dimensions, including the H plane tapers at each end. Table I gives the slot widths and depths studied. Based upon the Marcatili equations for λ_g, and also by considerations which will be given in the next section, the value d = 0.130" was chosen. This slot spacing puts the beam reasonably close to broadside for the frequencies of interest.

As a result of the radiation loss the propagation wavenumber now becomes complex, i.e.,

$$k_z = 2\ \pi/\lambda_g - i\ \alpha \ , \tag{6}$$

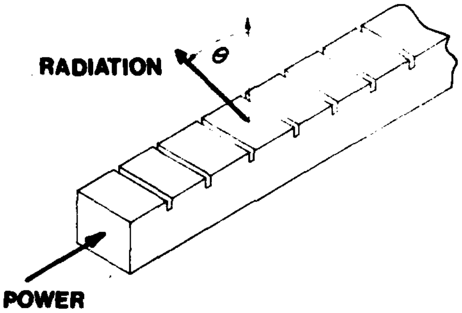


FIGURE 3. DIRECTION OF MAIN BEAM

	w (in)	t (in)
A	0.004	0.006
B	0.004	0.009
C	0.004	0.012
D	0.004	0.015
E	0.004	0.018
F	0.004	0.013
G	0.013	0.013
H	0.020	0.013

TABLE I. SLOT PARAMETERS

where α is the attenuation due to radiation along the length of the array, with $\alpha/k_z \ll 1$. The aperture excitation is therefore exponentially decaying, with the power attenuated equal to the power radiated, P_{RAD}, as given by

$$P_{RAD} = P_o \exp(-2\ \alpha\ L), \qquad (7)$$

where P_o is the power incident on the first slot, and L is the array length.

The Experiments

One waveguide was machined to size, as described above, from an unslotted slab. One tapered end was inserted into standard Ka band metal waveguide, using relatively low dielectric constant ($\epsilon_r = 2.2$) Rexolite as a spacer and support. Due to the large difference in the dielectric constants, the dielectric antenna characteristics were not effected by the Rexolite. The intrinsic insertion loss of the unslotted waveguide was measured, using the set-up shown in Figure 4. The measured loss was found to be 0.5 dB across the 30-38 GHz frequency range.

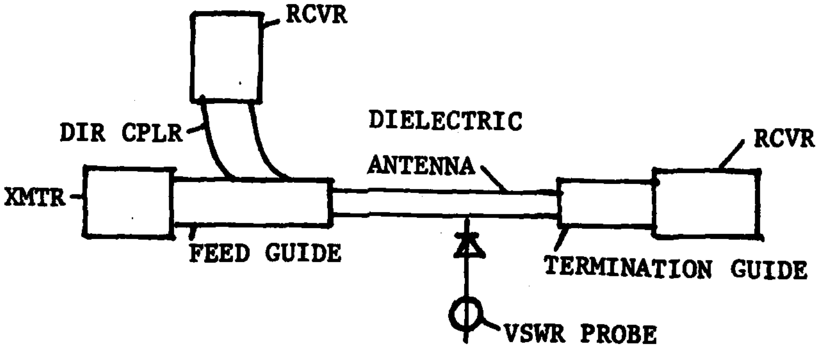


FIGURE 4. TEST SET-UP

Next, a reflecting short was attached to the other taper, and λ_g measured by external probe. These results appear as the upper curve in Figure 2.

Finally, an absorbing load was attached to the other end, so as to absorb any residue fields. Return loss was measured at the input end, and was better than 20 dB across the band, indicating that the combined H plane taper, guide cross section and load were all well matched.

Each slotted antenna was then similarly fitted and tested with the exception of the λ_g probe measurement, due to interference from the near-field radiation caused by the slots. Figure 5 gives some typical results. It is noted that there are loss resonances near 32 GHz and 36 GHz. The higher frequency resonance corresponds to the high reflection loss that occurs when the wavelength in the guide equals the slot spacing. For all antennas tested, it was found that the frequency at the center of the reflection loss curve corresponds to values of λ_g that lie between the upper and lower curves of Figure 2. In general, it was found that the width of the resonance, as measured 3 dB down from the peak of the reflection loss, did not exceed 1 GHz. As the reflection loss at 32 GHz is small ($\leq$ 1%), the lower frequency resonance corresponds to maximum power radiated by the antenna.

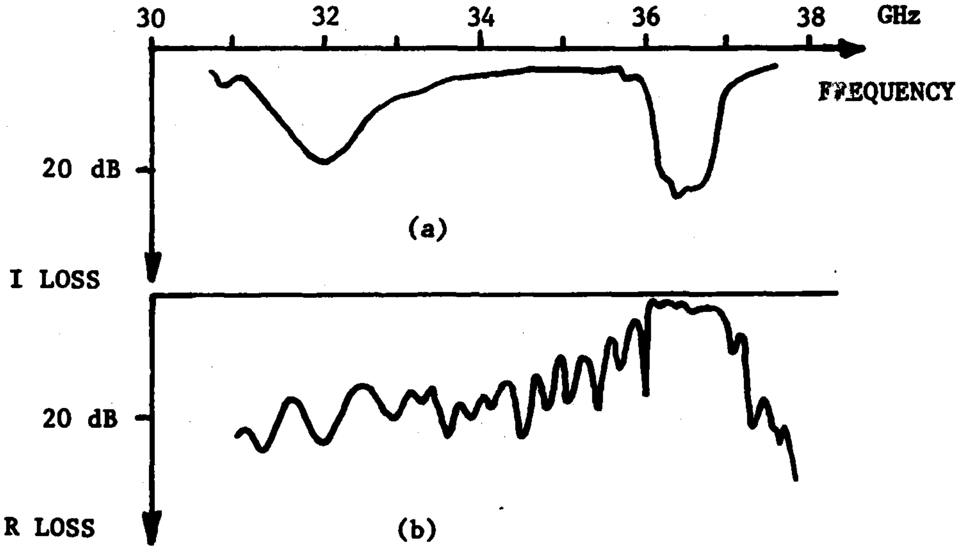


FIGURE 5. TYPICAL LOSS CURVES

(a) INSERTION LOSS
(b) RETURN LOSS

At any given frequency the ratio of net power loss ($P_{NET\ LOSS}$) (i.e., power not delivered to the load) to incident power is, from Figure 5(a)

$$\frac{P_{NET\ LOSS}}{P_o} = 1 - 10^{-I/10}, \tag{8}$$

where I is measured in dB. Taking into account the intrinsic 0.5 dB broadband loss, the power loss (P_{LOSS}) ratio becomes

$$\frac{P_{LOSS}}{P_o} = 1 - 10^{-I/10} - (1 - 10^{-0.05})$$

$$= 0.89 - 10^{-I/10} \tag{9}$$

Finally, the effect of power reflected back to the generator by the slots must be taken into consideration, so that the radiation loss ratio is

$$\frac{P_{RAD}}{P_o} = 0.89 - 10^{-I/10} - 10^{-R/10} , \quad (10)$$

where R is the return loss measured in dB from Figure 5(b). Combining Eq. (10) with Eq. (7) we find

$$\alpha = -\frac{1}{2L} \ln \left(0.89 - 10^{-I/10} - 10^{-R/10}\right). \quad (11)$$

Values of α obtained in this way are shown in Figure 6. The curves for antennas C, D, E and F are essentially identical. This appears to be the result of a saturation in the slot-surface wave coupling, as the slot depth is increased.

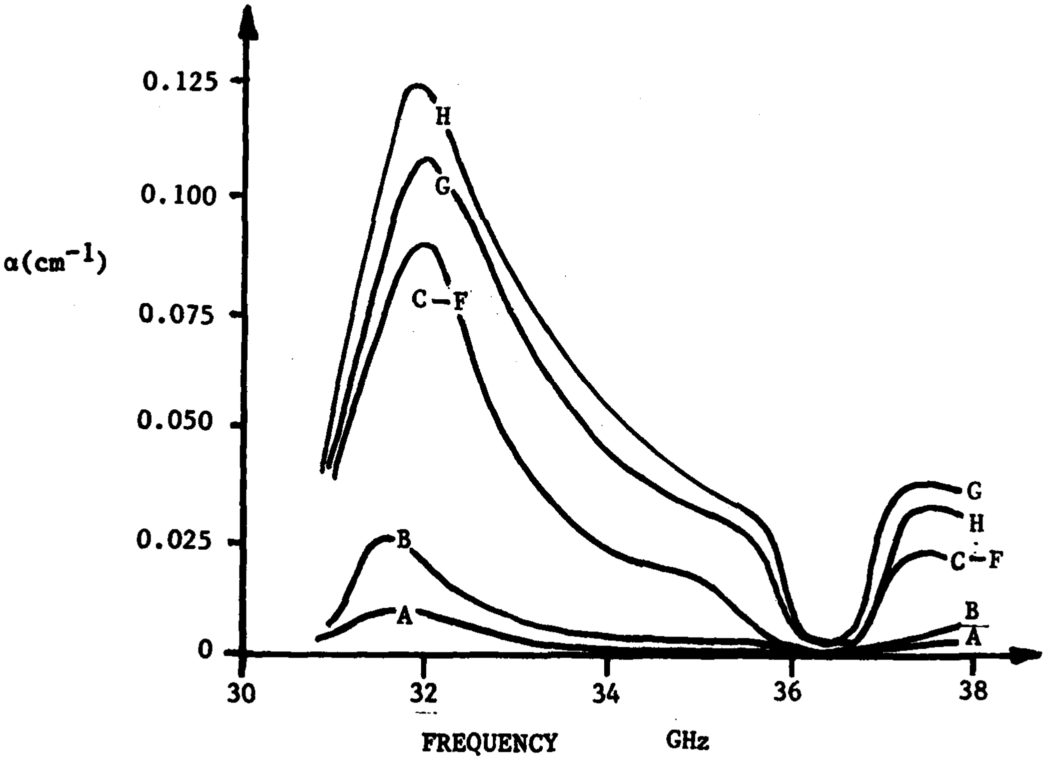


FIGURE 6. SLOT ATTENUATION VS. FREQUENCY

Far field radiation patterns were measured in an anechoic chamber. Several typical E-plane patterns are shown in Figure 7 at 32 GHz, and at 35 and 36 GHz, i.e., either side of broadside. It is noted that the first side lobe levels are typically 12 dB down, close to the theoretical value of 13 dB expected for an exponential illumination. The value of the expected angle of the main beam from the normal is given by Eq. (4), but we have worked this equation backwards, using the observed value of θ, to determine λ_g as a further check. In all cases we again have found λ_g to lie between the upper and lower curves of Figure 2. The 3 dB beamwidth is typically 6° over the 32-37 GHz band, but increases to about 8° near backfire at 31 GHz. The cross-polarization component was measured and found to be 20 dB down in the main beam.

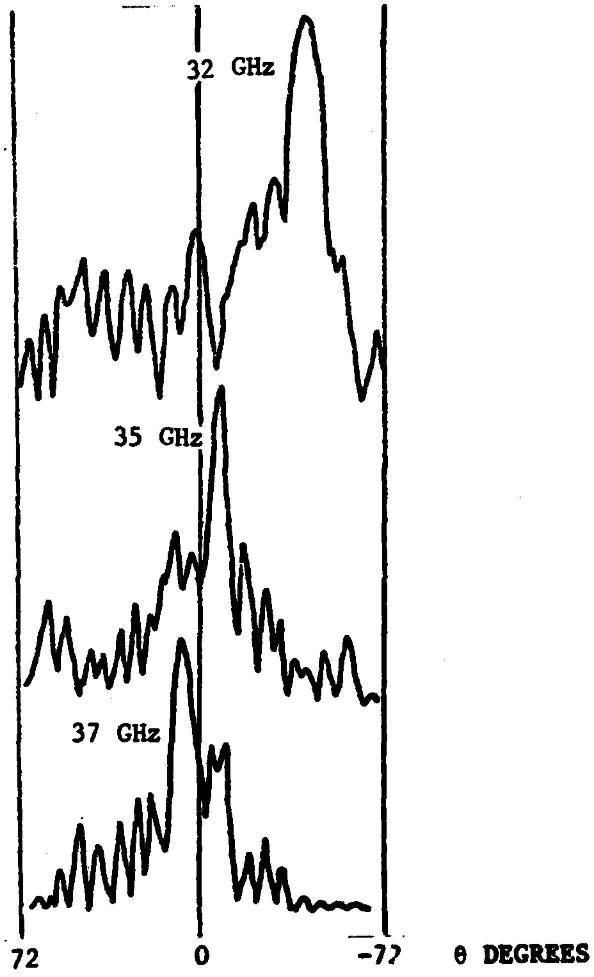


FIGURE 7. TYPICAL PATTERNS

The frequency scan characteristic is given by Figure 8; we find for all slot configurations studied a constant scan sensitivity of 11°/GHz, or approximately 4°/(percent change in frequency).

Gain was measured with respect to a standard gain horn. These results are shown in Figure 9.

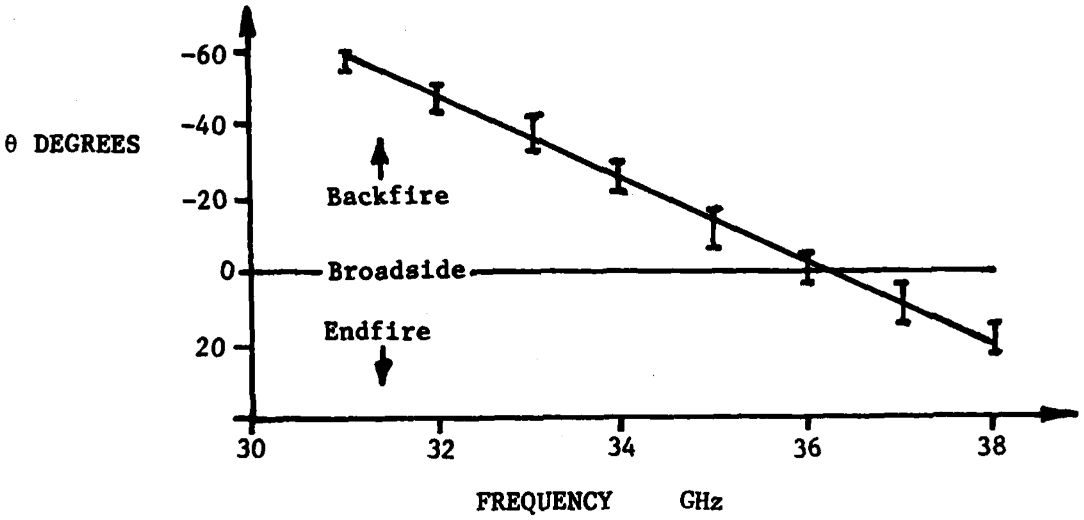


FIGURE 8. SCAN ANGLE VS. FREQUENCY

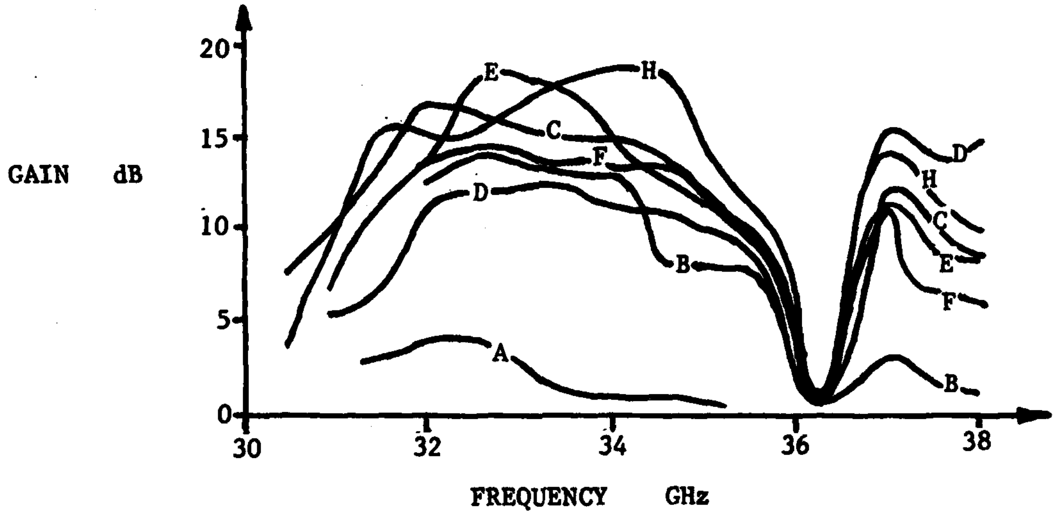


FIGURE 9. GAIN VS. FREQUENCY

Conclusions

Investigations have continued into developing design criteria for a dielectric waveguide line source antenna. The Marcatili analysis has proven to be a reliable guide for the transmission line design. There is, however, no practical design information available permitting the synthesis of a symmetrical amplitude taper along a grooved dielectric waveguide antenna. This investigation presents a measurement procedure that will permit a correlation between the radiating element physical configuration and the radiation properties. Additional data for various slot geometries will continue to be accumulated. A theoretical and emperical relationship will then be made so as to further relate the slot geometry to radiation characteristics. Our investigations have demonstrated that a non-metallic grooved dielectric waveguide can be made to radiate with a well-formed antenna pattern resulting from an exponential amplitude taper. This effort will be extended to permit the design of a lower sidelobe dielectric waveguide slot antenna.

Acknowledgement

The authors wish to thank Mr. William Johnson and Mr. Herbert Daley for their technical assistance, and Mrs. Phyllis Prorok for preparing the manuscript.

References

1. John Borowick, Richard A. Stern and Richard W. Babbitt, "A Small Electronic Scan Angle Antenna for Millimeter Waves", ERADCOM Report DELCS-TR-81-2, September 1981.

2. E.A.J. Marcatili, "Dielectric Rectangular Waveguide and Directional Coupler for Integrated Optics", Bell System Technical Journal, Volume 48, No. 7, pp 2071-2102, September 1969.

3. Johannes Jacobsen, "Analytical, Numerical, and Experimental Investigation of Guided Waves on a Periodically Strip-Loaded Dielectric Slab", IEEE Transactions on Antennas and Propagation, Volume AP-18, No. 3, pp 379-388, May 1970.

MICOM PROGRAM IN LIGHT CONTROL BY LIGHT (U)

CHARLES M. BOWDEN, DR.
RESEARCH DIRECTORATE, US ARMY MISSILE LABORATORY
US ARMY MISSILE COMMAND
REDSTONE ARSENAL, ALABAMA 35898

1. INTRODUCTION

Superfluorescence (1) is the phenomenon whereby a collection of atoms or molecules is prepared initially in a state of complete inversion, and then allowed to undergo relaxation by collective, spontaneous decay. This produces pulses of anomalous intensity (proportional to the square of the density of active atoms) and of anomalous, temporal duration (inversely proportional to the density). Since Dicke's initial work (2), there has been a preponderance of theoretical and experimental work dealing with this process (3).

With the exception of the more recent work of Bowden and Sung (4), all theoretical treatments have dealt exclusively with the relaxation process from a prepared state of complete inversion in a two-level manifold of atomic energy levels, and thus do not consider the dynamical effects of the pumping process. Yet, all reported experimental work (3,5-10) has utilized optical pumping on a minimum manifold of three atomic or molecular energy levels by laser pulse injection into the nonlinear medium, which subsequently fluorescesces.

It was pointed out by Bowden and Sung (4) that for a system otherwise satisfying the conditions for superfluorescent emission, unless the characteristic superradiance time (1), τ_R, is much greater than the pump pulse temporal duration, τ_p, i.e., $\tau_R >> \tau_p$, the process of coherent optical pumping on a three-level system can have dramatic effects on the SF. This is a condition which has not been realized over the full range of reported data.

In this paper, we present calculational results and analysis for the effects of coherent pump dynamics, propagation, transverse and diffraction effects on SF emission from an optically-pumped three-level system. The full, nonlinear, co-propagational aspects of the injected pump pulse, together with the SF which evolves are explicitly treated in the calculation.

Not only do our results relate strongly to previous calculations and experimental results in SF, but we introduce and demonstrate a new concept in nonlinear light-matter interactions, which we call light control by light. We show how characteristics of the SF can be controlled by specifying certain characteristics of the injection pulse. This leads to possible applications in laser pulse control and shaping, nonlinear optical information encoding, and unique pulsed LADAR and millimeter wavelength sources.

In the next section, the model upon which the calculation is based is presented and the algorithm used in the simulation is outlined. Results of the calculation are presented and discussed in Section 3. The last section is used to summarize the results and cite implications and to discuss future work.

2. MODEL FOR THREE-LEVEL SUPERFLUORESCENCE

The model upon which the calculation is based is comprised of a collection of identical three-level atoms, each having the energy level scheme shown in Figure 1. The 1 ↔ 3 transition is induced by a coherent electromagnetic field injection pulse of frequency ω_0 nearly tuned to the indicated transition. The properties of this pumping pulse are specified initially in terms of the initial and boundary conditions. The transition 3 ↔ 2 evolves by spontaneous emission at frequency ω . It is assumed that the energy level spacing is such that $\varepsilon_3 > \varepsilon_2 >> \varepsilon_1$ so that the fields of frequencies ω_0 and ω can be treated by separate wave equations. The energy levels 2 ↔ 1 are not coupled radiatively due to parity considerations.

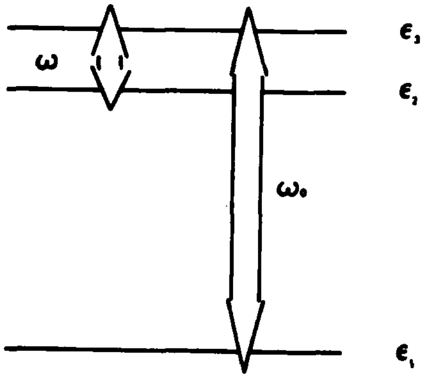


Figure 1. Model three-level atomic system and electric field tunings under consideration. For the results reported here, the injected pulse is tuned to the 1 ↔ 3 transition.

Further, we neglect spontaneous relaxation in the 3 ↔ 1 transition, and spontaneous relaxation in the 3 ↔ 2 transition is simulated by the choice of a small, but nonzero, initial transverse polarization characterized by the parameter $\phi_0 \sim 10^{-3}$. Our results do not depend upon nominal variations of this parameter. The initial condition is chosen consistent with the particular choice of ϕ_0, with nearly all the population in the ground state, and the initial values of the other atomic variables chosen consistently (4,11).

We use the electric dipole and rotating wave approximation and couple the atomic dipole moments to classical field amplitudes which are determined from Maxwell's equations. The Hamiltonian which describes the field-matter interaction for this system comprising N atoms (4), is

$$H = \hbar \sum_{r=1}^{3} \sum_{j=1}^{N} \varepsilon_{rj} R_{rr}^{(j)} - \frac{i\hbar}{2} \sum_{j=1}^{N} \left[\Omega^{(j)} R_{32}^{(j)} e^{-i(\omega t - \underline{k}\cdot\underline{r}_j)} - \Omega^{(j)*} R_{23}^{(j)} e^{i(\omega t - \underline{k}\cdot\underline{r}_j)} \right]$$

$$- \frac{i\hbar}{2} \sum_{j=1}^{N} \left[\omega_R^{(j)} R_{31}^{(j)} e^{-i(\omega_0 t - \underline{k}_0\cdot\underline{r}_j)} - \omega_R^{(j)*} R_{13}^{(j)} e^{i(\omega t - \underline{k}_0\cdot\underline{r}_j)} \right] \quad (1)$$

The first term on the right-hand side of Eq. (1) is the free atomic system Hamiltonian, with atomic level spacings ε_{rj}, r = 1,2,3; j=1,2,...,N. The second term on the right-hand side describes the interaction of the atomic system with the fluorescence field associated with the 3 ↔ 2 transition, whereas the last term on the right in Eq.(1) describes the interaction between the atomic system and the coherent pumping field. The fluorescence field and the pumping field have amplitudes $\Omega^{(j)}$ and $\omega_R{}^{(j)}$, respectively, in terms of Rabi frequency, at the position of the jth atom, $\underline{r}_j$. The respective wave vectors of the two fields are k and k_0 and the carrier frequencies are ω and ω_0. It is assumed that the electromagnetic field amplitudes vary insignificantly over the atomic dimensions and that all of the atoms remain fixed during the time frame of the dynamical evolution of the system.

The atomic variables in Eq. (1) are the canonical operators (4) $R_{k\ell}^{(j)}$ which obey the Lie algebra defined by the commutation rules (12-14)

$$\left[R_{ij}^{(m)} , R_{\ell k}^{(n)} \right] = R_{ik}^{(m)} \delta_{\ell j} \delta_{mn} - R_{\ell j}^{(m)} \delta_{ik} \delta_{mn} \quad (2)$$

i,j = 1,2,3; m,n = 1,2,...,N. The Rabi rates, $\Omega^{(j)}$ and $\omega_R^{(j)}$ are given in terms of the electric field amplitudes $E^{(j)}$ and $E_0^{(j)}$, respectively, and the matrix elements of the transition dipole moments, $u_{32}^{(j)}$ and $u_{31}^{(j)}$ by

$$\Omega^{(j)} = \frac{E^{(j)} u_{32}^{(j)}}{\hbar} \quad ; \quad \omega_R^{(j)} = \frac{E_o^{(j)} u_{31}^{(j)}}{\hbar} \quad , \tag{3a,b}$$

where we have considered only one linear polarization for the two fields and propagation in the positive z direction.

It is convenient to introduce a new set of variables in terms of the old ones. We let

$$W_{k\ell} = R_{kk} - R_{\ell\ell}\ , \ k > \ell \quad ; \quad R_{k\ell} = \frac{1}{2}(U_{k\ell} + i\ V_{k\ell})\ , \ k > \ell , \tag{4a,b}$$

where $U_{k\ell}, V_{k\ell}$, and $W_{k\ell}$ are real variables, and $U_{k\ell} = U_{\ell k}$, $V_{k\ell} = V_{\ell k}$,

$$\Omega = X + iY \quad ; \quad \omega_R = X_o + iY_o \quad , \tag{4c,d}$$

where X, Y, X_0 and Y_0 are real variables which are functions of both space and time.

The equations of motion for the atomic variables are calculated according to the commutation relation

$$i\hbar\ \dot{R}_{k\ell}^{(j)} = \left[H, R_{k\ell}^{(j)}\right] \quad . \tag{5}$$

If the transformation, Eqs. (4) is applied to the hierarchy of Eqs. (5), the resulting equations of motion for the real variables $\{W_{k\ell}, U_{k\ell}, V_{k\ell}\}$ are

$$\dot{W}_{31} = \frac{1}{2}\left\{XU_{32} - YV_{32}\right\} + \left\{X_oU_{31} - Y_oV_{31}\right\} - \gamma_{11}\left[W_{31} - W_{31}^{(e)}\right] , \tag{6a}$$

$$\dot{W}_{32} = \left\{XU_{32} - YV_{32}\right\} + \frac{1}{2}\left\{X_oU_{31} - Y_oV_{31}\right\} - \gamma_{11}\left[W_{32} - W_{32}^{(e)}\right] , \tag{6b}$$

$$\dot{U}_{32} = -\ \delta V_{32} - XW_{32} + \frac{1}{2}\left[X_oU_{12} - Y_oV_{12}\right] - \gamma_1 U_{32} , \tag{6c}$$

$$\dot{V}_{32} = \delta U_{32} + YW_{32} - \frac{1}{2}\left[X_oV_{21} + Y_oU_{21}\right] - \gamma_1 V_{32} , \tag{6d}$$

$$\dot{U}_{31} = \Delta V_{31} + \frac{1}{2}\left[XU_{21} + YV_{21}\right] - X_oW_{31} - \gamma_1 U_{31} , \tag{6e}$$

$$\dot{V}_{31} = -\ \Delta U_{31} + \frac{1}{2}\left[XV_{21} - YU_{21}\right] + Y_oW_{31} - \gamma_1 V_{31} , \tag{6f}$$

$$\dot{U}_{21} = \delta V_{21} - \frac{1}{2}\left[XU_{31} - YV_{31}\right] - \frac{1}{2}\left[X_o U_{32} - Y_o V_{32}\right] - \gamma_{\perp} U_{21} \ , \qquad (6g)$$

$$\dot{V}_{21} = -\delta U_{21} - \frac{1}{2}\left[XV_{31} + YU_{31}\right] - \frac{1}{2}\left[X_o V_{32} + Y_o U_{32}\right] - \gamma_{\perp} V_{21} \ , \qquad (6h)$$

where

$$\Delta^{(j)} = \epsilon_{33}^{(j)} - \omega_o \ , \quad \delta^{(j)} = \epsilon_{22}^{(j)} + \omega - \omega_o \ , \ \epsilon_{11} = 0. \qquad (7)$$

In obtaining Eqs. (6), we have made use of the invariant, tr S = I,

$$I \equiv R_{11}^{(j)} + R_{22}^{(j)} + R_{33}^{(j)} \ . \qquad (8)$$

It is noted that $\dot{I} = 0$ is satisfied identically in Eqs. (6) for $\gamma_{11} \to 0$. For $\gamma_{11} \neq 0$, the condition, Eq. (8) together with Eqs. (6) constitutes the statement of conservation of atomic density, i.e., particle number.

The Eqs. (6) are coupled to Maxwell's equations through the polarizations associated with each transition field. It is easily determined that the Maxwell's equations in dimensionless form in the rotating wave and slowly-varying envelope approximations can be written in the following form

$$\mathcal{F}_p^{-1}\nabla_\rho^2 \begin{Bmatrix} -\tilde{X}_o \\ \tilde{Y}_o \end{Bmatrix} + \frac{\partial}{\partial \eta_p} \begin{Bmatrix} \tilde{Y}_o \\ \tilde{X}_o \end{Bmatrix} = d \begin{Bmatrix} -U_{31} \\ V_{31} \end{Bmatrix} \ , \qquad (9a)$$

$$\mathcal{F}_s^{-1}\nabla_\rho^2 \begin{Bmatrix} -\tilde{X} \\ \tilde{Y} \end{Bmatrix} + \frac{\partial}{\partial \eta_s} \begin{Bmatrix} \tilde{Y} \\ \tilde{X} \end{Bmatrix} = d \begin{Bmatrix} -U_{32} \\ V_{32} \end{Bmatrix} \ , \qquad (9b)$$

where the variables $\tilde{X}$, $\tilde{Y}$, $\tilde{X}_o$, $\tilde{Y}_o$ are the same as those defined in Eqs. (4), but in units of $\gamma_{\perp}$. In the above equations, we have assumed cylindrical symmetry, thus

$$\nabla_\rho^2 = \frac{1}{\rho}\frac{\partial}{\partial \rho}\left(\rho \frac{\partial}{\partial \rho}\right) \ .$$

The first term on the left-hand side in Eqs. (9a,b) accounts for transverse effects with normalized radial coordinate $\rho = r/r_p$ where r is the radial distance and r_p is a characteristic spatial width. In Eqs. (9), $\eta_{p_s} = z \, g_{\mathrm{eff}_{p_s}}$ where $g_{\mathrm{eff}_{p_s}}$ is the on-axis effective gain,

$$g_{eff_{p_s}} = \frac{\begin{Bmatrix}\omega_o\\ \omega\end{Bmatrix}\begin{Bmatrix}\mu_{31}\\ \mu_{32}\end{Bmatrix}^2 N}{n\hbar c} T_2 \qquad (10)$$

where N is the atomic number density (assumed longitudinally homogeneous) and η is the index of refraction assumed identical for each transition wavelength. The quantity

$$d = N(r)/N_o \qquad (11)$$

governs the relative radial population density distribution for active atoms. Equations (9) are written in the retarded time, τ, frame where $\tau = t - n\, z/c$. From this point on, $\cdot$ in Eqs. (6) is taken to be $\cdot = \partial/\partial\tau$. Finally, the first factors on the first terms in Eqs. (9) are the reciprocals of the "gain length" Fresnel numbers defined by

$$\mathcal{F}_{p_s} \equiv \frac{2\pi r_p^2}{\lambda_{p_s}\, g^{-1}_{eff_{p_s}}} \quad . \qquad (12)$$

It is seen from Eqs. (9) that for sufficiently large Fresnel number, $\mathcal{F}$, the corrections due to transverse effects become negligible. The "gain length" Fresnel numbers $\mathcal{F}$ are related to the usual Fresnel numbers $F = 2\pi r_p^2/\lambda L$, where L is the length of the medium, by

$$\mathcal{F}/F = g_{eff}L \quad , \qquad (13)$$

i.e., the total gains of the medium. In the computations, diffraction is explicitly taken into account by the boundary condition that $\rho = \rho_{max}$ corresponds to completely absorbing walls.

The initial conditions are chosen to establish a small, but nonzero transverse polarization for the 3 ↔ 2 transition with almost the entire population in the ground state. This requires the specification of two small dimensionless parameters, $\varepsilon \sim 10^{-3}$, for the ground state initial population deficit, and $\delta \sim 10^{-3}$ for the tipping angle for the initial transverse polarization for the 3 ↔ 2 transition. The derivation for the initial values for the various matrix elements is presented elsewhere (15).

3. CALCULATIONAL RESULTS AND ANALYSIS

Calculational methods developed earlier (16) and discussed elsewhere (17,18) were applied to the model presented in the last section to compute the effects on SF pulse evolution for various initial conditions for the injected (pump) pulse. The results presented here demonstrate many facets of the control and shaping of the SF signal by control of the input signal initial characteristics.

In Figure 2 is shown results of the numerical calculation for the transverse integrated intensity profiles for the co-propagating SF and injected pulses at a penetration depth of z = 5.3 cm in the nonlinear medium. These profiles correspond to what would be observed with a wide aperture, fast, energy detector. The pumping pulses are labeled by capital letters and the corresponding SF pulses are labeled by the corresponding lower case letters. Each set of curves represents a different initial on-axis area for the pump pulse, i.e., curve A) is the reshaped pump pulse at z = 5.3 cm which had its initial on-axis area specified as $\theta_p = \pi$, and curve a) is the resulting SF pulse which has evolved. All other parameters are identical for each set of pulses. The initial conditions for the atomic medium is that nearly all the population is in the ground state ε_1 at $\tau = 0$, and a small, but nonzero macroscopic polarization exists between levels ε_3 and ε_2. These two conditions are specified by two parameters ε and δ, respectively, and we have chosen $\delta = \varepsilon = 10^{-3}$ self-consistently in terms of the initial values of the matrix elements for the entire system. These initial conditions are uniform for the atomic medium and are the same for all results reported here. Notice that we have neglected spontaneous relaxation in the pump transition, 1 ↔ 3, relative to the SF transition, 3 ↔ 2. This is justified due to our choice of relative oscillator strengths (see Figure 2 caption).

These results clearly indicate the coherence effect of the initial pump pulse area on the SF signal which evolves. Notice that the peak intensity of the SF pulses increases monatomically with initial on-axis area for the pump pulse. This is caused by self-focusing due to transverse coupling and propagation. For instance, a 2 π-injection pulse would generate a very small SF response compared to an initial π-injection pulse for these conditions at relatively small penetration z, or for the corresponding case in one spacial dimension. The delay time τ_D between the pump pulse peak and the corresponding SF pulse peak is very nearly inversely proportional to the input pulse area. The temporal SF pulse full width at half maximum (FWHM), τ_w is approximately invariant with respect to the injection pulse area.

Figure 3 shows the effect upon the SF pulse of variation in the initial temporal width at half maximum intensity for the pumping pulse. As the initial temporal width of the injected pulse τ_W becomes smaller, the SF delay time τ_D increases whereas the peak SF intensity decreases and the SF temporal width τ_s remains very closely fixed. It is clear from these results that there exists a linear relationship between the time delay τ_D,

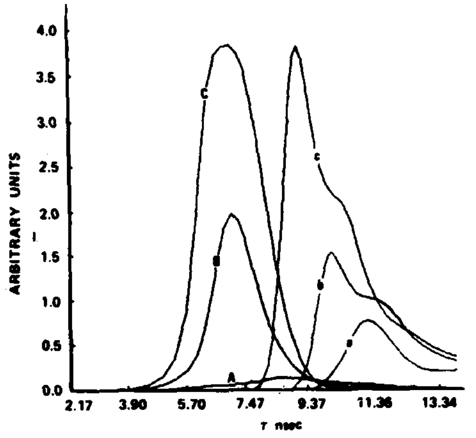


Figure 2. Radially integrated normalized intensity profiles for the SF and injected pulse at $z = 5.3$ cm penetration depth for 3 different values for the initial on-axis injection pulse area θ_p. The SF pulses are indicated by a, b, and c, whereas the corresponding injected pump pulses are labeled by A, B, and C. The injected pulses are initially Gaussian in r and τ with widths (FWHM) $r_0 = 0.24$ cm and $\tau_p = 4$ nsec, respectively. The level spacings are such that $(\varepsilon_3-\varepsilon_1)/(\varepsilon_3-\varepsilon_2) = 126.6$. The effective gain for the pump transition $g_p = 17\ \text{cm}^{-1}$ and that for the SF transition, $g_s = 291.7\ \text{cm}^{-1}$. The gain length Fresnel numbers for the two transitions are $\mathcal{F}_p = 16800$ and $\mathcal{F}_s = 2278$. The relaxation and dephasing times are taken as identical for all transitions, and are given as $T_1 = 80$ nsec and $T_2 = 70$ nsec, respectively. The injected pulse initial on-axis areas are: A) $\theta_p = \pi$; B) $\theta_p = 2\pi$; C) $\theta_p = 3\pi$.

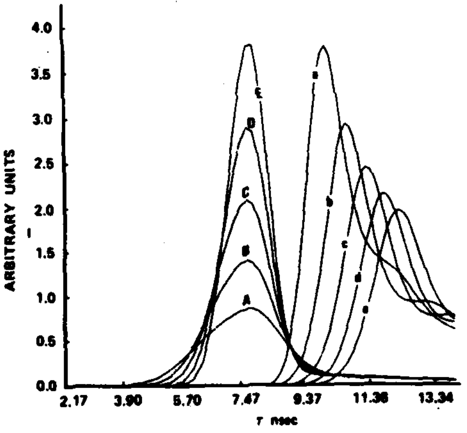


Figure 3. Radially integrated normalized intensity profiles for the SF and injected pulses at $z = 5.3$ cm penetration depth for 5 different values for the initial temporal width of the injected pulse. The initial on-axis area of the injected pulse is $\theta_p = \pi$ and the pump transition and SF effective gains are $g_p = 17.5\ \text{cm}^{-1}$ and $g_s = 641.7\ \text{cm}^{-1}$, respectively. All other parameters except for the Fresnel numbers, are the same as those for Figure 2. The injected pulse initial temporal widths at half maximum are: A) $\tau_w = 4$ nsec; B) $\tau_w = 3.3$ nsec; C) $\tau_w = 2.9$ nsec; D) $\tau_w = 2.5$ nsec; E) $\tau_w = 2.2$ nsec.

between the peak SF intensity and the corresponding pump pulse intensity, and the initial temporal width τ_w of the pump pulse. This is in qualitative agreement with the analytical prediction made in reference 4b, Eq. 5.1, based upon mean field theory. These results emphasize the importance of the initiating pulse characteristics in SF pulse evolution, and the effect of SF pulse narrowing with approximate pulse shape invariance by increasing the initial temporal width of the injected pulse. It is emphasized that all other parameters, including the initial value for the injected pulse on-axis area, are identical among these sets of curves.

The initial radial width, r_p, of the injected pulse was varied and the effect upon the SF pulse evolution is shown in Figure 4. There is clearly indicated an optimum value for r_p for which the SF peak intensity is a maximum and the SF temporal width τ_s is a minimum. If the relation, Eq. (13) is used in conjunction with the values of the parameters given in Figure 4 and its caption, it is seen that optimization occurs for a value for the conventional Fresnel number F_s for the SF transition, $F_s \simeq 1$. Thus, from Eq. (13) and $F_s = 1$, we have

$$\mathcal{F}_s = g_s z_{max} \tag{14}$$

for the gain-length Fresnel number. Since $F_s \sim 1/z$ the implication is that Eq. (14) gives the penetration depth, z_{max}, at which the SF peak intensity reaches a maximum in terms of the ratio $\mathcal{F}_s/g_s$. Since this takes both transverse and diffraction explicitly into account as well as propagation, this is indeed a profound statement.

Further insight into the implication of Eq. (14) can be obtained by considering a one-spacial dimension analogy. If the linear field loss is taken to be entirely due to diffraction, then the one-dimensional linear loss κ_s corresponding to the two-dimensional case specified by $\mathcal{F}_s$, is given by

$$\kappa_s = \frac{\lambda_s}{2\pi r_p^2} \quad . \tag{15}$$

Then, from Eq. (12)

$$\mathcal{F}_s = g_s/\kappa_s \tag{16}$$

is the effective gain, g_s, to loss, κ_s, ratio. From the condition, Eq. (14),

$$z_{max} = (\kappa_s^{-1}) \quad , \tag{17}$$

i.e., z_{max} is the penetration depth at which the SF peak intensity is a maximum, and corresponds to one effective diffraction length, as

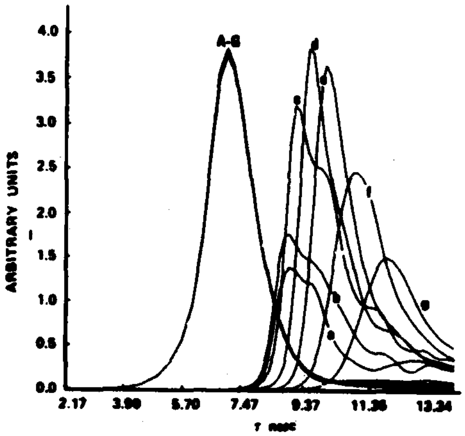


Figure 4. Radially integrated normalized intensity profiles for the SF and injected pulses at z = 5.3 cm penetration depth for 7 different values for the injected pulse initial radial width at half maximum r_p. The initial on-axis area θ_p of the injection pulse is $\theta_p = 2\pi$; the SF effective gain $g_s = 758.3\ \mathrm{cm}^{-1}$ and the pump transition effective gain $g_p = 14.6\ \mathrm{cm}^{-1}$; all other parameters are the same as for Figure 2. The initial radial widths at half maximum for the injected pulses are: a) $r_p = 0.57$ cm; b) $r_p = 0.43$ cm; c) $r_p = 0.24$ cm; d) $r_p = 0.18$ cm; e) $r_p = 0.15$ cm; f) $r_p = 0.11$ cm; g) $r_p = 0.09$ cm. The corresponding "gain length" Fresnel numbers are: a) $\mathcal{F}_p = 82876$, $\mathcal{F}_s = 34010$; b) $\mathcal{F}_p = 46898$, $\mathcal{F}_s = 19244$; c) $\mathcal{F}_p = 14428$, $\mathcal{F}_s = 5922$; d) $\mathcal{F}_p = 8326$, $\mathcal{F}_s = 3416$; e) $\mathcal{F}_p = 5602$, $\mathcal{F}_s = 2298$; f) $\mathcal{F}_p = 3422$, $\mathcal{F}_s = 1404$; g) $\mathcal{F}_p = 2082$, $\mathcal{F}_s = 854$.

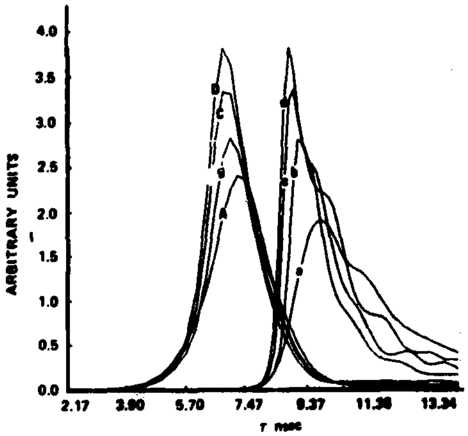


Figure 5. Radially integrated normalized intensity profiles for the SF and injected pulses at z = 5.3 cm penetration depth for 4 different values for the injected pulse initial radial shape parameter ν (see text). The initial on-axis area θ_p of the injected pulse is $\theta_p = 2\pi$ and the SF effective gain $g_s = 758.3\ \mathrm{cm}^{-1}$, whereas the effective gain for the pump transition $g_p = 14.6\ \mathrm{cm}^{-1}$. All other parameters are the same as for Figure 2. The initial radial shape parameters for the injected pulses are: A) $\nu = 1$; B) $\nu = 2$; C) $\nu = 3$; D) $\nu = 4$.

defined by Eq. (15). Carrying the one-dimensional analogy one step further, Eq. (16) used in Eq. (13) gives

$$F = (\kappa z)^{-1} \quad . \tag{18}$$

From Eqs. (16) and (18) we have exhibited the significance of the Fresnel numbers $\mathcal{F}$ and F in terms of diffraction loss, i.e., $\mathcal{F}$ can be thought of as gain to loss ratio, Eq. (16), whereas F can correspondingly be thought of as the reciprocal of the strength of the diffraction loss, Eq. (18).

The effect on SF pulse evolution of variation of the initial radial shape of the initiating pulse is shown in Figure 5. The shape parameter ν is defined in terms of the initial condition for the pump transition field amplitude, $\omega_R(r)$,

$$\omega_R(r) \;=\; \omega_R(o) \exp\left[- (r/r_p)^{\nu}\right] . \tag{19}$$

Thus, for $\nu = 2$, the initial amplitude of the injected pulse is radially Gaussian, whereas for $\nu = 4$, it is radially super-Gaussian. We see from the results presented in Figure 5 that as the initial radial shape of the injected pulse becomes broader, i.e., larger values for ν , the peak intensity of the SF pulse generated becomes larger, and the width τ_S diminishes. It is emphasized that all other parameters, including the initial values for the radial and temporal widths are invariant among these sets of curves.

The response of SF pulse evolution to changes in the initial temporal shape of the injection pulse is shown in Figure 6 which compares the effect of a Gaussian initial temporal shape for the pump pulse, identified by the temporal shape parameter, $\sigma = 2$, with that of a super-Gaussian identified by $\sigma = 4$. As for the radial distribution discussed previously, the temporal shape parameter σ is defined in terms of the initial condition for the pump transition field amplitude $\omega_R(\tau)$,

$$\omega_R(\tau) \;=\; \omega_R(o) \exp\left[- (\tau/\tau_p)^{\sigma}\right] . \tag{20}$$

Again, it is seen that the broader initial pump pulse causes an increase in the peak SF intensity and a reduction in the delay time τ_D and SF pulse width τ_s.

Finally, Figure 7 shows the effect of variation of the density ρ of active atoms. The effective gains, g_S and g_P, are changed proportionally, corresponding to a density variation ρ . The ratio of the SF intensities is $I_c/I_b = 1.76$ and $I_b/I_a = 2.06$; these ratios are larger than the

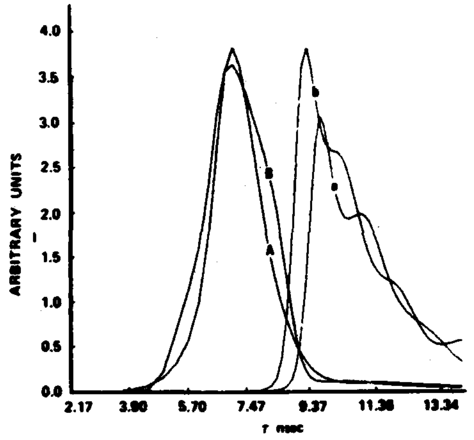


Figure 6. Radially integrated normalized intensity profiles for the SF and injected pulses at $z = 5.3$ cm penetration depth for two different values for the injected pulse initial temporal shape parameter σ (see text). The initial on-axis area θ_p of the injected pulse is $\theta_p = 2\pi$, and the SF effective gain $g_s = 641.7$ cm^{-1}. All other parameters are the same as for Figure 4(c). The initial radial shape parameters for the injected pulses are: a) $\sigma = 2$; b) $\sigma = 4$.

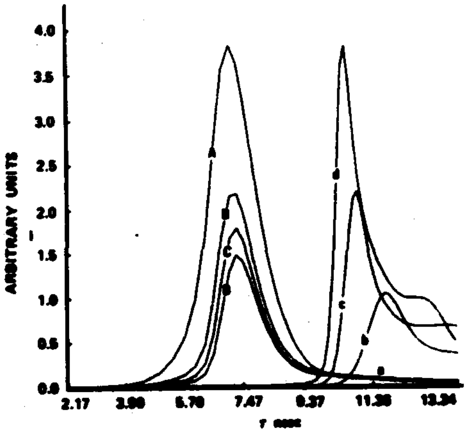


Figure 7. Radially integrated normalized intensity profiles for the SF and injected pulses at $z = 5.3$ cm penetration depth for three different values for the density ρ of atoms. The on-axis initial area θ_p for the injected pulse is $\theta_p = 2\pi$. Except for the effective gains and Fresnel numbers, the values for all other parameters are the same as for Figure 4(c). For each set of curves, the gain values are: a) $g_s = 525.0$ cm^{-1}; $g_p = 26.3$ cm^{-1}; b) $g_s = 641.7$ cm^{-1}, $g_p = 32.1$ cm^{-1}; c) $g_s = 758.3$ cm^{-1}, $g_p = 37.9$ cm^{-1}. The corresponding Fresnel numbers are: a) $\mathcal{F}_p = 25992$, $\mathcal{F}_s = 4100$; b) $\mathcal{F}_p = 31724$, $\mathcal{F}_s = 5010$; c) $\mathcal{F}_p = 37465$, $\mathcal{F}_s = 5922$.

corresponding density ratios squared (1), $(\rho_c/\rho_b)^2 = 1.40$ and $(\rho_b/\rho_a)^2 = 1.49$. This difference may be due to self-focusing, especially since the values of the effective gains used in this case are quite high. However, the ratio of the temporal widths, τ_W, (FWHM), is within 15% of the corresponding inverse ratios of the densities; the same is true for the delay time, τ_D, of the SF intensity peak with respect to the pump intensity peak. These results compare qualitatively reasonably well with the mean field predictions for SF in two-level systems initially prepared in a state of complete inversion (1).

4. CONCLUSIONS

The results presented here clearly demonstrate the coherence and deterministic effects on SF pulse evolution of injection pump pulse initial characteristics and conditions. It is suggested that effects of the type discussed here may have in fact been operative in SF experiments and their results which were published earlier (5-10).

Furthermore, and perhaps of greater importance, we have demonstrated the control and shaping of the SF pulse which evolves, by specification of particular initial characteristics and conditions for the pumping pulse which is injected into the nonlinear medium to initiate SF emission. These manifestations and others of the same class, we call the control of light by light via a nonlinear medium. This phenomenon constitutes a method for nonlinear information encoding, or information transfer, from the injection pulse characteristics to corresponding SF pulse characteristics which evolve due to propagation and interaction in the nonlinear medium.

Work is now in progress to incorporate the effects of quantum statistics of the SF spontaneous relaxation process (19). We are in the process of further determination and analysis of the nonlinear interaction between two co-propagating pulses resonantly, as well as nonresonantly, interacting by a nonlinear medium (19).

ACKNOWLEDGEMENT

Continued and valued collaboration with F. P. Mattar, Y. Claude and M. Cormier is gratefully acknowledged.

REFERENCES

1. R. Bonifacio and L. A. Lugiato, Phys. Rev. A11, 1507 (1975); A12, 587 (1975).

2. R. H. Dicke, Phys. Rev. 93, 99 (1954).

3. See papers and references in Cooperative Effects in Matter and Radiation, edited by C. M. Bowden, D. W. Howgate and H. R. Robl, Plenum, New York, (1977). "Swept-gain Superradiance in CO_2-pumped CH_3F", J. J. Ehrlich, C. M. Bowden, S. H. Lehnigk, A. T. Rosenberger and T. A. DeTemple, in Coherence and Quantum Optics IV, edited by L. Mandel and E. Wolf, Plenum, New York, p. 923 (1978). "MIRADCOM Program in Swept-gain Super-Radiance", D. W. Howgate, C. M. Bowden and J. J. Ehrlich, US Army Science Conference Proceedings, June (1978), First Prize.

4. C. M. Bowden and C. C. Sung, Phys. Rev. A18, 1558 (1978); Phys. Rev. A20, 2033 (1979).

5. N. Skirbanowitz, J. P. Herman, J. C. MacGillivray and M. S. Feld, Phys. Rev. Lett. 30, 309 (1973).

6. H. M. Gibbs, Q. H. F. Vrehen and H. M. J. Hickspoors, Phys. Rev. Lett. 39, 547 (1977).

7. Q. H. F. Vrehen, in Cooperative Effects in Matter and Radiation, edited by C. M. Bowden, D. W. Howgate and H. R. Robl, Plenum, New York, p. 79 (1977).

8. M. Gross, C. Fabre, P. Pillet and S. Haroche, Phys. Rev. Lett. 36, 1035 (1976).

9. A. Flusberg, F. Mossberg and S. R. Hartmann, in Cooperative Effects in Matter and Radiation, edited by C. M. Bowden, D. W. Howgate and H. R. Robl, Plenum, New York, p. 37, (1977).

10. A. T. Rosenberger and T. A. DeTemple, Phys. Rev. A24, 868 (1981).

11. F. T. Hioe and J. H. Eberly, Phys. Rev. Lett. 47, 838 (1981).

12. R. Gilmore, Lie Groups, Lie Algebras, and Some of Their Applications, Wiley, New York, Chapter 6, Section 2, (1974).

13. R. Gilmore, C. M. Bowden and L. M. Narducci, Phys. Rev. A12, 1019 (1975).

14. R. Gilmore, C. M. Bowden and L. M. Narducci, "c-Number Representation for Multilevel Systems and the Quantum-Classical Correspondence", in Quantum Statistics and the Many-Body Problem, edited by S. B. Trickey, W. R. Kirk and J. W. Dufty, Plenum, New York, (1975).

15. F. P. Mattar and C. M. Bowden, "Coherent Pump Dynamics and Pulse Evolution in Three-level Superfluorescence and Control of Light by Light", in Topics in Current Physics: Multiple Photon Dissociation of Polyatomic Molecules, edited by C. D. Cantrell, Springer-Verlag, New York, to appear, (1982).

16. F. P. Mattar and M. C. Newstein, in Cooperative Effects in Matter and Radiation, edited by C. M. Bowden, D. W. Howgate and H. R. Robl, Plenum, New York, p. 139 (1977).

17. F. P. Mattar, in Optical Bistability, edited by C. M. Bowden, M. Ciftan and H. R. Robl, Plenum, New York, p. 503 (1981); in Proceedings, Tenth Simulation and Modeling Conference, Pittsburgh, 1978, edited by W. Vogt and M. Mickle, Publ. Inst. Soc. Am., Pittsburgh, PA (1979).

18. F. P. Mattar, H. M. Gibbs, S. L. McCall and M. S. Feld, Phys. Rev. Lett. 46, 1123 (1981).

19. C. M. Bowden and F. P. Mattar, to be published.

MICROBIAL ATTACK OF NITROCELLULOSE[a]

BRUCE W. BRODMAN and *MICHAEL P. DEVINE
U.S. ARMY ARMAMENT RESEARCH AND DEVELOPMENT COMMAND
DOVER, NEW JERSEY 07801

Deterrents are materials diffused some distance into nitrocellulose (NC)-based small arms propellant grains in order to slow down their burning rate early in the ballistic cycle when the propellant bed surface area is at maximum. A variety of deterrent materials are used in both ball and extruded military propellants. In the case of ball propellants, the specific deterrent utilized is di-n-butyl phthalate (DBP). A series of past studies(1-3) has revealed the concentration profile produced when DBP is diffused into a spherical propellant grain and established the mechanism involved. Propellant grains deterred in this manner have the disadvantage of lot-to-lot ballistic variations along with the possibility that deterrent migration might occur under extreme storage conditions. Further, the deterring manufacturing operation involves raising the temperature to about 70°C. If this high temperature exposure could be avoided, it would result in a safety improvement. For these reasons an alternate method for control of the gas generation schedule would be desirable.

The present study is directed toward the use of microorganisms to remove energetic nitrate ester groups from the surface of propellant grains in order to obtain the requisite control of the propellant burning rate.

Past work regarding the microbial decomposition of NC produced conflicting results. Bokorny(4) found that mold grew on NC suspended in an aqueous medium containing mineral salts. The suggestion was made that the NC provided the mold with essential carbon and perhaps nitrogen. Malenkovic(5) and Jacque(6) in independent investigations came to the conclusion that the organism was utilizing only dissolved mineral salts and not attacking the NC. The authors concluded that the organisms gave rise to organic compounds such as acids, which in turn can degrade the NC. More

[a] Parts of this presentation have previously appeared in the Journal of Applied Polymer Science, Vol. 26, 997-1000 (1981).

recently, studies(7) involved with the microbial decomposition of NC particles in waste water concluded that direct denitration of the NC did not occur; however, appropriate chemical pretreatment, such as with a mineral acid, accomplished hydrolysis of the nitrate ester. The free nitrate ions could then be utilized by the microorganism.

EXPERIMENTAL

Nitrocellulose. In all cases, ground pyroxylin (11.11% nitrogen) was used.

Culture medium. The initial culture medium consisted of KH_2PO_4, K_2HPO_4, and $MgSO_4$ each present at 0.7 g/liter. In addition, the solution contained trace amounts of sodium chloride, iron sulfate, zinc sulfate, and magnesium sulfate.

Inoculation and incubation. In all cases, the organism used was Aspergillus fumigatus. In each case, the culture medium was added to a 250-ml Ehrlenmeyer flask along with 1% by weight of ground pyroxylin. The flask was then incubated at 31°C at 200 rpm on a rotary shaker for the requisite time.

Mycelial weight. The mycelial weight was obtained by filtering the flask contents followed by repeated washing with an ether-alcohol mixture. This wash removed the NC and the isolated mycelia were dried and weighed.

DISCUSSION

No fungus growth was observed in the initial experiments wherein ground pyroxylin was suspended in the nitrogen deficient culture medium. Growth was observed, however, when 3% glucose was added to the shaking culture flask containing the pyroxylin, culture medium, and organism. Experiments were run in order to determine the weight of mycelial tissue produced by Aspergillus fumigatus in the glucose and pyroxylin containing medium. Table I contains the growth data. Examination of Table I indicates that essentially no growth occurred when the pyroxylin was absent. The fact that glucose is needed for growth would tend to indicate that the organism could not utilize the carbon from the NC. Further, the need for pyroxylin would indicate that the organism was utilizing the nitrogen from the NC.

An experiment was run to detrmine if the microbial attack of NC involved direct utilization of nitrogen or if the nitrate ester group had undergone hydrolysis prior to utilization by the organism. This type of hydrolysis has been previously reported in the literature.(8) Pyroxylin (1 g) was shaken in both a 100-ml flask of distilled water and in a 3% glucose-inorganic salt solution for seven days. The pyroxylin was filtered off, and to the flask containing only distilled water was added a 3% glucose-concentrated salt solution in order to adjust the salt concentration to

TABLE I

Mycelial Dry Weights of *Aspergillus fumigatus* Grown on Glucose-Salts Medium Containing Ground Pyroxylin as Sole Source of Nitrogen

Incubation time, days	Mycelial dry weights, mg, nitrogen source None	 pyroxylin, 1%
5	0.7, 1.7	39.7, 34.9, 37.6
11	4.6	92.5, 49.8
21	0.6, 1.4	59.0, 63.8, 60.4
40	2.3, 2.3	77.2, 73.4, 74.3

that used in the previous experiment. All filtrates were cold sterilized by membrane filtration and then inoculated with a spore suspension of *Aspergillus fumigatus*. In addition, a flask containing a 3% glucose solution and an inorganic nitrogen source (0.67 g of $NaNO_3$) at a level equivalent to 1 g of pyroxylin (111 mg of N) was inoculated with *Aspergillus fumigatus*.

The flasks were incubated at 31°C for 21 days on a rotary shaker. Weights of mycelia produced with the various conditions are listed in Table II. It should be noted that the mycelia were not washed with ether-alcohol prior to weighing, and thus were higher than washed weights. Examination of Table II indicates that both the distilled-water extract and salt extract contained sufficient nitrogen to support growth. Interestingly enough, the extract made with the glucose-salt solution appeared to extract slightly more nitrogen than did the distilled-water extract. This may have been due to the lower pH of the salt extract which enhanced the hydrolysis rate of the nitrate ester group. As would be expected, the cultures containing the inorganic nitrogen source showed significantly more growth. Also, a culture containing no nitrogen or carbon sources showed essentially no growth.

The next phase of the study involved an effort to establish the rate of hydrolysis of the NC nitrate ester group when suspended in an aqueous medium. Milled pyroxylin (1 g) was suspended in 100 ml of distilled water and shaken on a rotary shaker. At the appropriate intervals, the pyroxylin was separated from the water by centrifugation and the pyroxylin resuspended in fresh distilled water. The nitrate ion concentration of the supernatant liquid was determined by means of an Orion specific ion electrode and an Orion 801 ion analyzer. Table III summarizes the results of the experiment.

Chemical analysis of the dried mycelia of *Aspergillus fumigatus* has shown a nitrogen content of 2%. The values listed in Table III for nitrate ion show that about 0.44 mg of nitrogen can be removed 1 g of nitrocellulose as nitrate ion, and that 95% of this is removed in the first four days. Based on the 2% N content of Aspergillus fumigatus and an available nitrogen content of 0.44 mg, the mycelial mass one would expect would be about 22.0 mg. This value of 22.0 mg for the mycelial mass is lower than the experimental value obtained for five days of growth (Table I).

Table I gives data indicating greater growth than could be accounted for by the experimental hydrolysis data. This fact can be accounted for by enhanced nitrocellulose nitrate ester hydrolysis caused by the organism.

TABLE II

Mycelial Dry Weights Produced during Growth of Aspergillus fumigatus for 21 Days on Solution of Glucose and Mineral Salts with $NaNO_3$ or on Extracts of Ground Pyroxylin as Nitrogen Source

Carbon source	Nitrogen source	Mycelial dry weight, mg replicates	average
None	None	5.3, 5.7, 5.5	5.5
3% glucose	0.67 g $NaNO_3$	634.6, 652.3, 656.6	647.8
3% glucose	pyroxylin extract[a]	22.8, 49.4, 48.1	40.1
3% glucose	pyroxylin extract[b]	33.4, 30.6, 34.3	32.8

[a] Shaken for seven days at 200 rpm at 31°C in solution containing glucose, 30.0 g/liter; $MgSO_4 7H_2O$, 0.7 g/liter; K_2HPO_4, 0.7 g/liter; KH_2PO_4, 0.7 g/liter; NaCl, 0.005 g/liter; $FeSO_4 7H_2O$, 0.002 g/liter; $ZnSO_4 7H_2O$, 0.002 g/liter; $MnSO_4 H_2O$, 0.001 g/liter.

[b] Shaken for seven days as above except with distilled water alone. Salts and glucose were added to distilled-water extract before cold sterilization of complete medium.

TABLE III

Nitrate Levels in 100 ml Distilled-Water Extracts of 1 g Pyroxylin

Extraction time, days	Average,[a] NO_3^-	Average[b] mg N/flask	Replicates
4, initial extraction	2.90×10^{-4}M	0.406	4
6, second extraction	1.49×10^{-5}M	0.021	6
4, third extraction	0.74×10^{-5}M	0.011	6

[a] Distilled water alone gave (NO_3^-) readings of less than 1×10^{-5}M.

[b] 1 g pyroxylin contains 111.1 mg nitrogen.

In summary, it has been shown that Aspergillus fumigatus can utilize nitrogen from pyroxylin suspended in a nitrogen deficient medium if a supplementary carbon source is provided. It appears that the organism does not directly attack the nitrocellulose but rather utilizes the nitrogen resulting from the hydrolysis of the nitrocellulose nitrate ester group. Further, it appears that the organism caused enhanced hydrolysis of the nitrocellulose.

REFERENCES

1. B. W. Brodman, M. P. Devine, and M. T. Gurbarg, J. Appl. Polym. Sci., 18, 943 (1974).
2. B. W. Brodman, M. P. Devine, R. W. Finch, and M. S. MacClaren, J. Appl. Polym. Sci., 18, 3739 (1974).
3. B. W. Brodman, J. A. Sipia, Jr., and S. Schwartz, J. Appl. Polym. Sci., 19, 1905 (1975).
4. T. Bokorny, (as quoted by T. Urbanski) in Chemistry and Technology of Explosives, Vol. 2, Pergamon, London, England, 1965, p. 313.
5. B. Malenkovic, (as quoted by T. Urbanski) in Chemistry and Technology of Explosives, Vol. 2, Pergamon, London, England, 1965, p. 313.
6. M. Jacque, (as quoted by T. Urbanski) in Chemistry and Technology of Explosives, Vol. 2, Pergamon, London, 1965, p. 313.
7. T. M. Wendt and A. M. Kaplan, Treating Waster Water Containing Cellulose Nitrate Particles, U.S. Pat. 3,939,068 (Feb. 17, 1976).
8. J. Desmaroux, (as quoted by T. Urbanski) in Chemistry and Technology of Explosives, Vol. 2, Pergamon, London, England, 1965, p. 397.

AN EMPIRICAL MODEL FOR NEAR MILLIMETER WAVE SNOW EXTINCTION AND BACKSCATTER (U)

*Douglas R. Brown, Ph.D.

US Army Atmospheric Sciences Laboratory
White Sands Missile Range, New Mexico 88002

A snow extinction and backscatter model for near millimeter waves was developed to complete a set of natural atmosphere propagation models for inclusion in the Electro-Optics System Effects Library (EOSAEL) (1). This model is compact, simple, fast, and uses commonly available meteorological data as inputs. The accuracy is commensurate with the accuracy of those input parameters; where the input parameters do not properly characterize the environment, further input specification is available. For snow, the rain equivalent accumulation rate is the fundamental input, with air temperature providing the distinction between dry and wet snowtypes. For millimeter wave (MMW), the classification of the snowflake by its ice-to-liquid-water ratio is necessary, due to the large differences in the complex indices of refraction of ice and water (2).

The snow extinction model is presented first, including a review of the data used to generate it. The backscatter model is then discussed, followed by a comparison between the model and measurements made at the Cold Regions Research and Engineering Laboratory (CRREL) SNOW-ONE test in January-February 1982 (3).

1. SNOW EXTINCTION

While the literature on the interaction of MMW with snow is limited, the complete literature will not be reviewed here; a general review has been carried out by Kobayashi (4). Only those results used in the model will be discussed.

The general classification of snowtypes is discussed in detail by Nishitsuji (5). Four types are described: dry, moist, wet, and watery, distinguished by the density of the snowflake. Because no other determination of snowflake density was reported in the literature, it is difficult to follow this classification procedure; in practice, snowtypes were inferred by comparison with the Nishitsuji extinction results where possible. Data for extinction by wet snow at 15 GHz, moist snow at 35 GHz, and watery snow at 50 GHz are presented. A Mie scattering type model was developed by Nishitsuji, but is cumbersome, and only his experimental results were employed.

Data at 11, 15, 24, and 48 GHz are presented by Oomori and Aoyagi (6), although no indication of the snowtypes is given. A unique confocal resonator was used to measure snow attenuations. Comparison with other data seems to indicate that moist snow propagation was measured. The data are modeled with an aR^b relation, which is derived using a theory of snowfall statistics and in which R is rain equivalent snowrate and a and b are parameters. The a parameter data agrees well with other data, but at 24 and 48 GHz there is poor agreement for the b parameter.

Moist snow data is provided at 35 GHz by Robinson (7). The data set is of limited extent, having only seven useable values. Malinkin et al (8) measured dry snow at 35 GHz, with but a slight improvement: eight points.

Only two data points exist in the literature for frequencies above 50 GHz: one at 140 GHz for wet snow by Richard, Kammerer, and Reitz (9), and one at 312.5 GHz for dry snow by Babkin et al (10). The Richard paper also reports visual range for the same snow data. Readers should note the error in reference (9), figure 25, where the a parameter is written as 1.37 but clearly should be 3.7 instead.

Analysis of the data shows that an aR^b relation for snow is the simplest form for a snow model, but does such a model have theoretical foundations? A review of aR^b models for rain extinction provides the guidelines.

Relations of the form aR^b can be derived for rain in the Rayleigh and optical limits. This result relies on the use of negative exponential or modified gamma dropsize distribution and a power series representation for the particle forward scattering amplitude. Where such power series representations cannot be found, direct fits to Mie scattering calculations are used. Such models provide highly accurate results for only a limited range of rainrates, but application over wider ranges is useful because of the general dominance of errors introduced from the beginning by the arbitrary raindrop size distribution. Errors can be significant, but may be minimized by the judicious selection of raindrop distribution (11).

Errors in the snow model can be expected to be even larger. Errors of drop or snowflake distribution are compounded by the snowflake orientation dependence of the forward scattering amplitude, which now is due to a highly asymmetric flake or aggregate. Some of this asymmetry will be reduced by volume averaging over the falling snowflake orientations, but large variations due to snowflake type must be anticipated. The similiarity or lack thereof between the a and b parameter frequency and particle size dependencies will indicate how complete this averaging process is. For snowflakes that are truly amorphous, with no orientation forces in free fall, this technique should provide a reasonable model. This cannot be expected to be true for all snowtypes. If flake orientations are not random, then significant fluctuations in extinction and backscatter from the model will occur. Because this model neglects polarization, it must also be used cautiously, since polarization effects can be expected to be large for snow. The treatment of polarization and snowflake orientation are beyond the scope of a model of this type. There are probably few, if any, models that can even attempt to treat these effects in a general manner for snow.

The a parameters taken or derived from the literature are shown in figure 1 along with the various distribution parameters used in a corresponding rain model. The scarcity of snow data is evident, but the data are consistent with expectation. Dry snow, representing pure ice crystals and air, has small MMW refractive indices relative to liquid water, and correspondingly low extinction. As snow becomes wetter, its scattering cross section increases dramatically--an effect seen as the radar "bright band." At nigher frequencies, as scattering approaches the optical limit, the cross section evidences a $^1/R$ dependence. Dry snow, with its predominance of small flakes, has less extinction at low frequencies, but as the optical limit is approached the relationship of dry and wet snow extinction will reverse, giving dry snow greater attenuation than wet snow. Where this crossover point occurs is not indicated in this data set. All the curves pass through the 312 GHz point for lack of data at these high frequencies.

The use of an aR^b relation for snow is supported by this data, which does not display large unexpected variations from the rain cases. Note that at low frequencies, below 35 GHz where the Rayleigh approximation applies, the curves for wet snow and rain do not converge, due to refractive index differences and to the fact that the extinction is expressed as a function of accumulation rates, not mass density.

The b parameter, which determines the dependence of extinction on accumulation rate, is shown in figure 2, again with rain values and the inferred frequency and snowflake behavior.

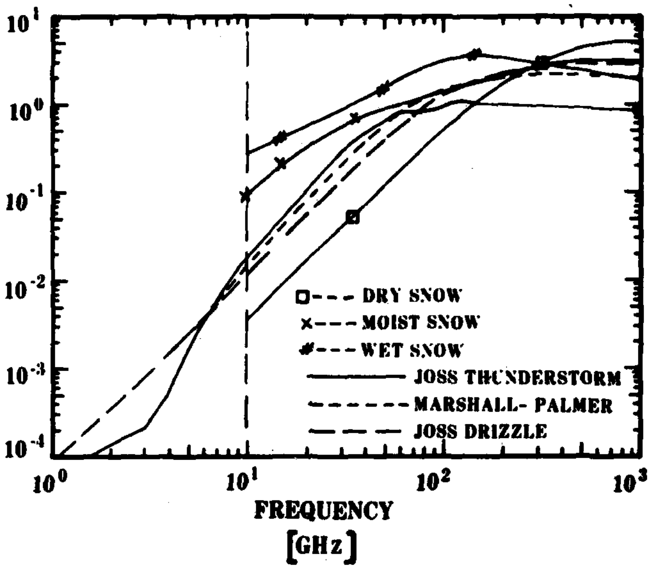


Figure 1. Snow parameter a with rain comparison.

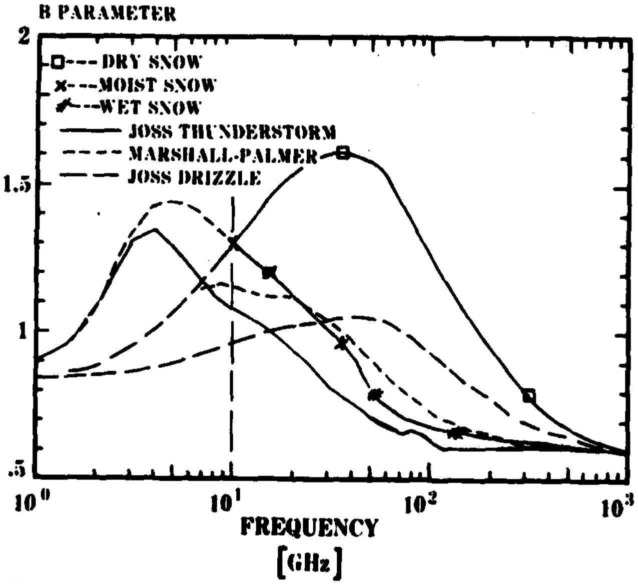


Figure 2. Snow parameter b with rain comparison. Data extrapolated below 10 GHz

The wet and moist snow data are similar to the Joss thunderstorm and Marshall-Palmer behavior; these curves have been roughly followed in defining the wet and moist snow b parameter curves. Dry snow behavior is expected to follow the Joss drizzle distribution due to the predominance of small particles in both cases. However, two points provide only the crudest indication of the dry snow behavior. The large peak value of the dry snow b parameter should not cause alarm, because dry snowrates are generally quite low, there being little or no data for rates above 2 mm/h in the literature.

Air temperature at the ground is used to determine the snowtype for an extinction calculation. Dry snow is used at or below freezing, moist snow parameters are used from 0°C to 2°C, and wet snow is employed above 2°C. These temperatures are rough estimates, based upon the assumption that temperature lapse rates are negative and are subject to adjustment depending on the specific storm characteristics.

The frequency range covered by this model is not well defined. With eight data points defining three curves and with only two measurements above 50 GHz, the high frequency behavior of the model is obviously subject to error. For this reason, use of the model above 100 GHz should be accompanied with caution.

2. SNOW BACKSCATTER

MMW snow backscatter data is severly limited by the use of inferred reflectivity factors derived from measured snowflake size distributions. Such results are valid only in the Rayleigh regime below 35 GHz. Initially, a Rayleigh model was employed with restricted frequency coverage. With the publication of a datum at 95 GHz resulting from the SNOW-ONE measurements, it was decided, due to the inadequacy of the existing model, to develop a new model using all the available data. No validation of this model is possible because all the data were incorporated in it.

The similar behavior of the rain and snow a and b parameters led to the use of the rain backscatter model to provide the shape of the snow backscatter curves. This procedure can be dangerous, but theory and experiment are not yet able to provide sufficient knowledge to define an independent model. At long wavelengths, the reflectivity factors of Imai et al (12) are used along with the Rayleigh relation to compute backscatter cross section, η,

$$\eta = \frac{\pi |k|^2}{\lambda^4} Z$$

where $k = (m^2 - 1)/(m^2 + 2)$, m is the complex index of refraction, λ the wavelength, and $Z = \int N(D)D^6 \, dD$ where N(D) is the snowflake size distribution function. In practice, reflectivity is related to rain equivalent snowrate as

$$Z = aR^b \ .$$

Table 1 lists the values of the a and b parameters from Imai et al, where their classification is assumed to correspond with our dry, moist, and wet classes.

TABLE 1. RAYLEIGH BACKSCATTER PARAMETERS

Snowtype	a	b
dry	600	1.8
moist	1800	1.8
wet	2400	1.8

At shorter wavelengths, the rain backscatter curve was forced to go through the datum from Nemarich et al (13) at 95 GHz. Their nine data points were abstracted into a single upper bound value. This was necessary because their data did not correlate well with the snow accumulation data, an apparent example of snowflake resonant scattering variations. The value used was 6.6×10^{-5} m^2/m^3 for a dry snowrate of 1 mm/h. Examples of the computed snow backscatter cross sections for the three snowtypes and a snowrate of 5 mm/h are shown in figure 3. Comparison with rain backscatter indicates that for dry snow the cross section is much less for snow; for moist snow, rain and snow are comparable; and for wet snow, snow backscatter is much larger than rain.

In scaling the snowtype results, it was assumed that the Imai et al results would hold at all frequencies, thus giving moist snow three times and wet snow four times the dry snow cross section. This may not be correct, but no data are available yet to improve the assumption.

3. MODEL EVALUATION

The SNOW-ONE test (3) of January and February 1981 provided the first opportunity to evaluate the snow extinction model using a totally independent data set.

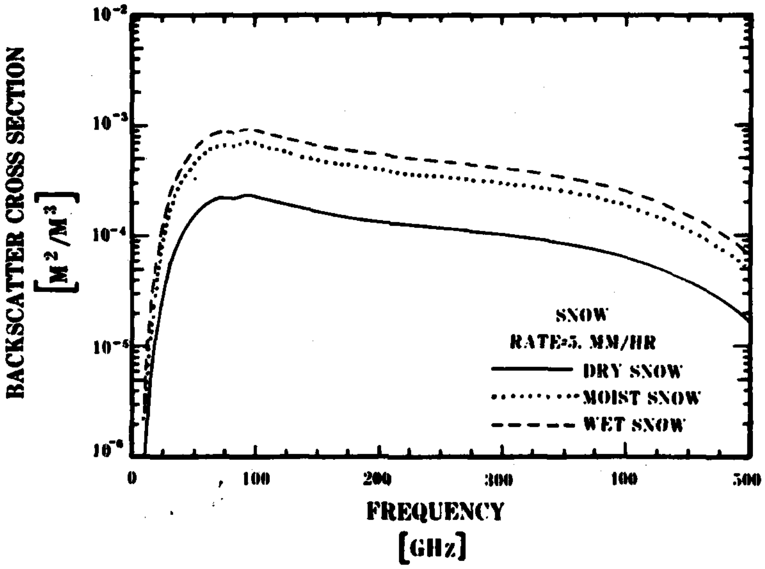


Figure 3. Snow backscatter cross section.

Both Harry Diamond Laboratories (HDL) and the US Army Ballistics Research Laboratory (BRL) carried out MMW propagation experiments at SNOW-ONE. The HDL Mobile Measurement Facility, operating at 95, 140, and 225 GHz, measured extinction and backscatter, reported by Nemarich et al (13). The BRL experiment measured extinction at 35, 95, 140, and 217 GHz, reported by Bauerle and Knox (14). Due to the small data sample at 35 GHz, this frequency is not included in the evaluation.

Supporting meteorological and snow characterization data have been provided by CRREL (3) and the US Army Atmospheric Sciences Laboratory (15).

The most significant storm of the test, beginning at 0400 on 8 February and finishing at 0500 on 9 February, was selected for analysis. Snow accumulation data was available from 1700 on the 8th to 0400 on the 9th. At 2200 on the 8th windspeeds increased, producing significant amounts of blowing snow and ending the capability of the snow accumulation gauges to indicate the amount of airborne snow. Therefore, only the data from 1700 to 2200 on 8 February are suitable for comparison with the model predictions. To provide a sense of the variations in extinction, the data for the period from 1700 on the 8th to 0400 on the 9th will be presented.

The millimeter transmission data was reported relative to the clear path transmission determined either before or after the storm. The measurements do not correct for temperature and relative humidity variations during the storm, but the relative humidity remained above 95 percent during this period, and temperature variations are estimated to introduce at most errors of ±0.06 dB/km at 95 GHz, ±0.15 dB/km at 140 GHz, and ±0.4 dB/km at 220 GHz. These are small enough to be neglected.

The presence or absence of fog during the storm can influence millimeter extinction. Light scattering instruments were used to measure aerosol size distributions, but their results are not easy to interpret. At the succeeding SNOW-ONE-A test held in Dec 81 and Jan 82, a background aerosol was measured at all times, presumably due to the wood stoves used to heat many houses in the region. At SNOW-ONE this background aerosol was not anticipated and therefore not measured. It is not possible to separate unequivocally the fog or smoke aerosol contribution to the aerosol distributions measured at SNOW-ONE. During the period from 1700 to 2200, equating all the measured aerosol with fog provides only a maximum of 0.035 gm/m^3 of fog liquid water, and generally much less. This corresponds to a maximum of 0.17 dB/km at 95 GHz, 0.24 dB/km at 140 GHz, and 0.29 dB/km at 220 GHz, all less than 10 percent of the extinctions measured during that time. Apparent fog densities did increase dramatically after 2200, but this was certainly due to the blowing snow.

Rather than compute the extinction from the measured snowrate, the measured snow extinction was used in an inverse calculation to predict the snow accumulation rate. This was done to provide the most concise temporal presentation of the three frequency results. Figures 4 and 5 show the "inverse" model predictions compared to the measured snowrate for the BRL and HDL data, respectively. Conversion from moist to dry snow at 1850, following the temperature decrease, has been made. For the period from 1700 to 2200 the agreement is quite good for an empirical model with such a limited data base. The peak at 217 GHz in the BRL data at 2100 is not repeated in the HDL data and cannot at this time be explained. The consistency of the model is very good, and it is tempting to conclude that the snow accumulation measuring equipment slightly undersamples. Such a conclusion is unwarranted, but does indicate that traditional snowrate measurements are really not adequate for the detailed characterization required by this data.

The large excursions after 2200 are due to the blowing snow, but the disagreement between the HDL and BRL results at 225 and 217 GHz may be due in part to improper calibration of the HDL 225 GHz data because the correct calibration was not yet available. Since this snow model cannot address the conditions after 2200, discussion is academic at this time.

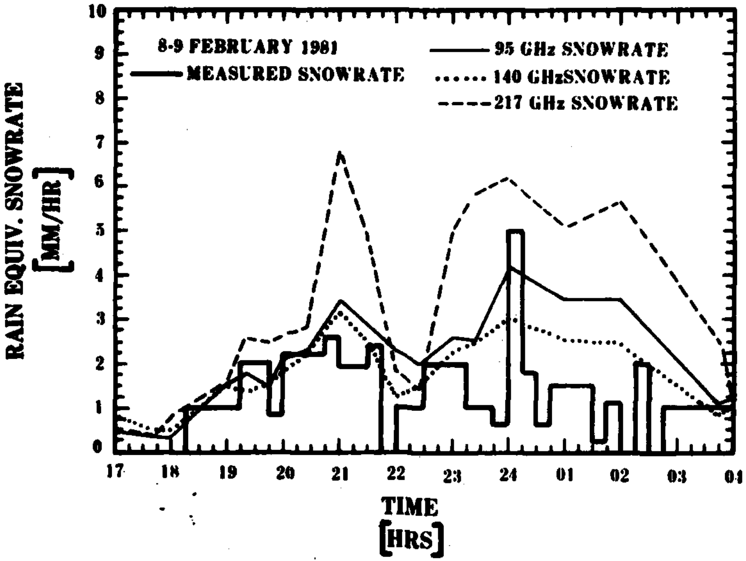


Figure 4. Snowrate computed from BRL data compared to measured snowrate.

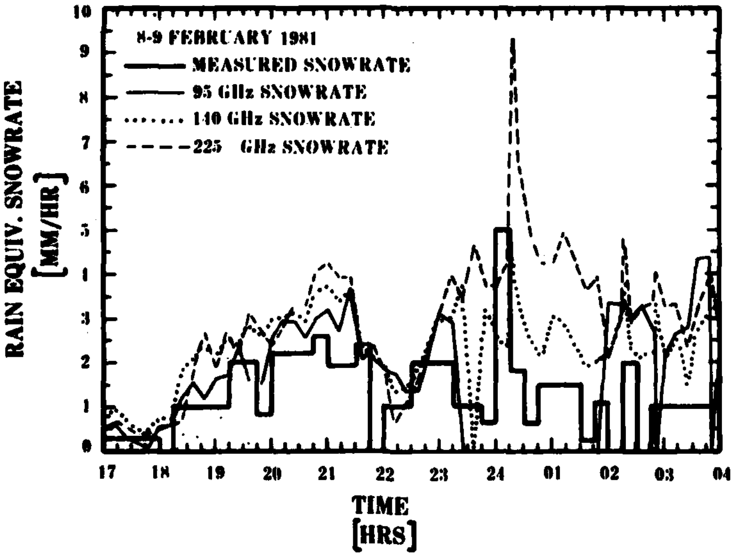


Figure 5. Snow rate computed from HDL data compared to measured snowrate.

These results give added confidence to the use of the model at frequencies up to 225 GHz, at least for dry snow. Wet and moist snow accuracy must still be determined at these high frequencies. Given existing field test programs, it is unlikely that measurements above 225 GHz will be available, so that extension of the model with confidence above this frequency may not be possible.

4. CONCLUSION

A new MMW snow model has been presented, based upon the use of an aR^b relation. The data used in its development were briefly reviewed, and the similarity of the snow and rain models was shown. The extinction model was compared with independent data from two separate facilities and demonstrated remarkably good agreement. The backscatter model could not be evaluated because no independent data exist.

The development of this model has clarified three data deficiencies: lack of snow backscatter data from all types of snow, lack of polarized propagation data at all frequencies, and lack of unpolarized data at frequencies above 100 GHz. Efforts to provide for the inadequacies of the data base continue, but snow data is difficult to obtain, and further development of the model will be slow.

A particular problem is that a large data base is necessary at each frequency to define the a and b parameters, due to the use of snowrate as the fundamental quantity. This also limits the model to near calm conditions, since there is no way to estimate the effects of blowing snow. A more sophisticated model using airborne snow density or the snowflake size distribution function itself could be developed, but the data set would be even smaller than that employed here. Since these data are not part of the standard meteorological data base, it would be most difficult to apply such models over various geographic and diverse climatological areas. For these reasons, this empirical MMW snow extinction and backscatter model is offered as a solution to the predictive problem of the effects of snow on MMW propagation.

5. REFERENCES

1. Duncan, Louis D., EOSAEL 80, Volume 1, Technical Documentation, ASL-TR-0072, US Army Atmospheric Sciences Laboratory, White Sands Missile Range, New Mexico, 1981.

2. Ray, P. S., "Broadband Complex Refractive Indices of Ice and Water," Applied Optics, 11:1836-1844, 1972.

3. Redfield, R. K., ed., SNOW-ONE Preliminary Data Report, Internal Report #715, US Army Cold Regions Research and Engineering Laboratory, Hanover, New Hampshire, 1981.

4. Kobayashi, Herbert K., Effect of Hail, Snow, and Melting Hydrometeors on Millimeter Radio Waves, ASL-TR-0092, US Army Atmospheric Sciences Laboratory, White Sands Missile Range, New Mexico, 1981.

5. Nishitsuji, A., "Method of Calculation of Radio-Wave Attenuation in Snowfall," Elect and Comm in Japan, 54-B:34, 1971.

6. Oomori, T., and S. Aoyagi, "A Presumptive Formula for Snowfall Attenuation of Radio Waves," Elect and Comm in Japan, 54-B:74, 1971.

7. Robinson, N. P., "Measurements of the Effect of Rain, Snow, and Fogs on 8.6 mm Radar Echoes," Proceedings IEEE, 203-B:209, 1955.

8. Mahlinkin, V. G., et al, "Attenuation of Signal at the Wavelength $\lambda = 8.6$ mm in Hydrometeors," Rad Eng and Elect Physics, 21:1, 1976.

9. Richard, W. W., J. E. Kammerer, and R. G. Reitz, "140-GHz Attenuation and Optical Visibility Measurements of Fog, Rain, and Snow," Memorandum Report ARBRL-MR-2800, US Army Ballistics Research Laboratory, Aberdeen Proving Ground, Maryland, 1977.

10. Babkin, Yu. S., et al, "Attenuation of Radiation at a Wavelength of 0.96 mm in Snow," Radio Eng and Elect Physics, 15:2171, 1970.

11. Keizer, W., J. Snieder, and C. de Haan, "Rain Attenuation Measurements at 94 GHz: Comparison of Theory and Experiment," DPP Symposium, Neubiberg bei Munchen, Germany, 1978.

12. Imai, I., M. Fujiwara, I. Ichimura, and Y. Toyama, "Radar Reflectivity of Falling Snow," Meteorology and Geophysics (Japan), 6:130, 1955.

13. Nemarich, J., R. J. Wellman, D. Rocha, and G. B. Wetzet, "Characteristics of Near-Millimeter Wave Propagation in Snow," SPIE, Vol 305, in press.

14. Bauerle, D., and J. E. Knox, "Millimetre Wavelength Radar Propagation and Backscatter Measurements at SNOW-ONE," Snow Symposium I, in press.

15. Olsen, R. O., D. R. Brown, J. E. Butterfield, and R. Orkasinski, "SNOW-ONE Atmospheric and Transmission Measurements," Snow Symposium I, in press.

AURORA FLASH X-RAY FACILITY AS A SOURCE-REGION EMP SIMULATOR

M. Bushell, R. Manriquez,
G. Merkel, and *W. D. Scharf
US Army Electronics Research and
Development Command
HARRY DIAMOND LABORATORIES
Adelphi, MD 20783

The study of source-region electromagnetic pulse phenomenology and coupling[1-4] is distinguished from more conventional electromagnetic research by the presence of time-varying air conductivity. The relevant Maxwell equations are:

$$\begin{aligned} &\varepsilon \nabla \cdot E = \rho \ , \\ &\mu \nabla \cdot H = 0 \ , \\ &\nabla x E = -\mu \dot{H} , \\ &\nabla x H = J + \sigma E + \varepsilon \dot{E} \end{aligned}$$

These can be combined into a wave equation for the electric field:

$$\nabla^2 E - \mu(\varepsilon \ddot{E} + \sigma \dot{E} + \dot{\sigma} E) = \mu \dot{J} + \nabla(\rho/\varepsilon) \quad .$$

(The magnetic field can be found by time-integrating $\nabla x E$.) The first two terms appear in the well-known homogeneous wave equation:

$$\nabla^2 E - \frac{1}{c^2} \frac{\partial^2 E}{\partial t^2} = 0 \ ,$$

where $c^2 = \frac{1}{\mu\varepsilon}$.

The third term represents the dissipation of energy resulting from air conductivity--i.e., the collision of charged free carriers (electrons and heavy ions) with neutral molecules. When this term dominates the second, the system becomes overdamped and is described by the diffusion equation:

$$\nabla^2 E - \mu\sigma \frac{\partial E}{\partial t} = 0 \quad .$$

The fourth term contributes only when σ is time-varying, and even then is significant only when the time variation is as fast or nearly as fast as that of $E(t)$. (Note that the third and fourth terms together represent $\frac{\partial}{\partial t}(\sigma E)$-- the time derivative of conduction current density.)

The right-hand side terms represent the electromagnetic sources. In the case of nuclear EMP, these sources are provided by high energy radiation from a nuclear device--through the mechanism of the Compton effect (which leads to space curent and charge separation)--and by current and charge density induced in local or distant objects (including the ground). The region in which $\vec{J}(\vec{x},t)$ and $\rho(\vec{x},t)$ are non-zero is called the source region, and the term "souce-region EMP" is used whenever these local drivers are significant. In a typical burst, the source region may be several miles in linear extent, and sources throughout the region contribute, through the well-known mechanisms of electromagnetic radiation and diffusion, to the fields at any given point.

This last observation lies at the heart of the greatest challenge to successful SREMP simulation. Powerful sources of high-energy pulsed radiation (e.g., AURORA[5,6], and Hermes[7,8]) are available, and these naturally produce Compton current and charge separation in the air (or other gas) into which the radiation is released. There is a temptation to assume that if these local sources are reasonably well reproduced, the full electromagnetic environment will be also. This naive assumption neglects the contribution of "distant" sources that cannot be reproduced, since existing pulsed-radiation technology is capable of irradiating a volume whose linear dimensions are expressed in tens of meters at most. (The AURORA test cell measures 20 m x 12 m x 5 m.)

Of course, due to the finite velocity of light, the "prompt" environment can be reliably reproduced. However, after 10 nanoseconds or so, the absence of sources at a distance of more than 20 feet from the field point becomes noticeable. In addition, if the simulation testing is performed in a volume enclosed by metal walls (as in AURORA), this effect is aggravated by the electromagnetic boundary conditions imposed by these walls. They short out the electric field. Figure 1 compares (not to scale, but schematically) the AURORA test cell volume to the vast distribution of Compton electron source currents in an actual SREMP. Figure 2 is a schematic view of the inside of the AURORA test cell during a shot.

A less subtle limitation of facilities like AURORA is that the radiation pulse itself does not faithfully reproduce the pulse produced by a nuclear weapon. It rises more slowly and decays more quickly. The gamma spectrum is not quite right. The late-time drivers are very different, largely due to the lack of neutrons.

In spite of these limitations, the authors, who have for several years been conducting an experimental program at AURORA,[9-27], have found that useful and relevant work can be done there. AURORA is a fine source of "basic physics" data on SREMP environments and coupling. Further, we believe that testing of the nature described here should be incorporated in any plan for SREMP vulnerability assessment and hardness validation of military and civilian electrical and electronic equipment. While ultimate determination of vulnerability and hardness will rely also on computer calculations, we feel that AURORA testing plays an essential role, both in validating the codes and in exposing unexpected vulnerabilities.

SREMP environments can, in a rough way, be described as belonging to one

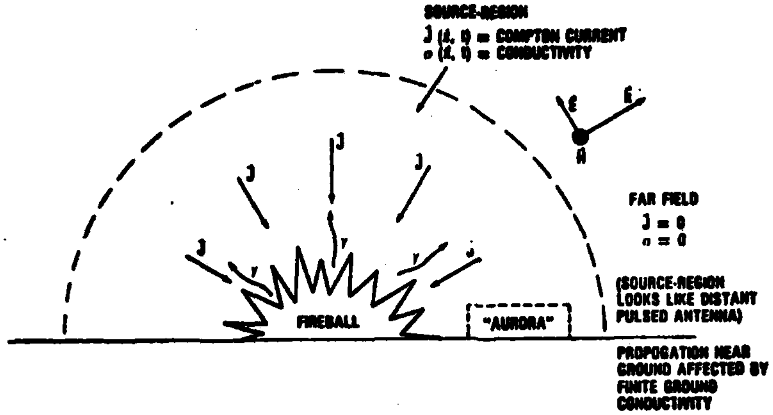

SOURCE-REGION
J (r, t) = COMPTON CURRENT
σ (r, t) = CONDUCTIVITY
E
E
H
J
J
J
J
J
γ
γ
γ
FAR FIELD
J = 0
σ = 0
(SOURCE-REGION LOOKS LIKE DISTANT PULSED ANTENNA)
FIREBALL
"AURORA"
PROPOGATION NEAR GROUND AFFECTED BY FINITE GROUND CONDUCTIVITY

Figure 1. Comparison of small volume of the AURORA test cell with the vast distribution (several miles diameter) of Compton current drivers in an actual SREMP. Actually, the relative size of the test cell is much smaller than shown in the figure.

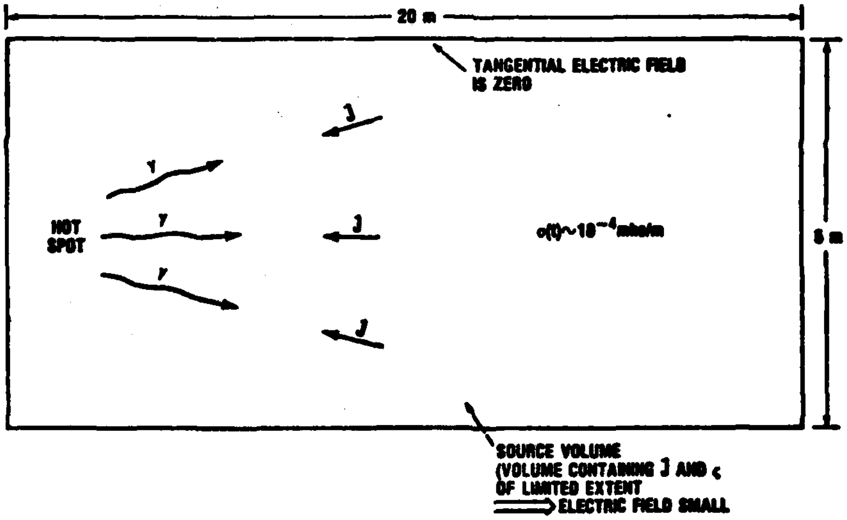

20 m
TANGENTIAL ELECTRIC FIELD IS ZERO
J
γ
HOT SPOT
γ
J
σ(t)~10⁻⁴mho/m
5 m
γ
J
SOURCE VOLUME (VOLUME CONTAINING J AND σ OF LIMITED EXTENT ⟹ ELECTRIC FIELD SMALL

Figure 2. Schematic view of the AURORA test cell during an AURORA shot. The metallic test cell walls can short out the electric field generated by the Compton electrons.

of two categories--tactical or strategic. These two types of environment, and several approaches to their simulation, will be discussed in turn.

The Tactical Case

Tactical source-region EMP is the electromagnetic component of a nuclear environment in which a soldier can reasonably be expected to remain alive and functioning for a useful time. Thus, tactical equipment, for use in a battlefield, is typically hardened only to a level consistent with a "man-survivability criterion." In practice, this normally means that peak conductivity levels of more than 10^{-3} mho/m are of no interest. (Long cables, which may channel energy from the "deep source region" out to a less highly dosed area, is an exception to this general rule.)

Some rough indication of the time evolution and magnitudes of relevant parameters in the tactical case is given by the graphs of figure 3. These are taken from calculations made using LEMP[28] (a highly respected environment code in general use by the EMP community) and assuming a set of device characteristics[29] which describe no existing weapon, but can be taken as typical.

The primary driver is, of course, the dose rate, $\dot{\gamma}$, expressed in units of rads/s. Thus the $\dot{\gamma}$ pulse shape determines the shapes of all relevant time histories--i.e., $\vec{J}(\vec{x},t)$, $\sigma(\vec{x},t)$, $\vec{E}(\vec{x},t)$, and $\vec{H}(\vec{x},t)$.

In early time, the most significant descriptive parameter is the 10%-90% risetime. This rise time, typically on the order of 10 nanoseconds, is critical to the efficiency of resonant coupling to relatively small objects, with linear dimensions of ten feet or so. Most Army equipment (such as radios, vehicles, shelter-resident fire control systems, etc.) falls into this category.

In late time, the pulse can be described by one or more exponential decay constants. These decay times are much longer than the risetime. At late time, the resonant behavior of "small" objects is not relevant, and such objects can be regarded with considerable accuracy as simple current collectors--driven directly by a combination of Compton and conduction current. (Displacement curent, important at early time, is of practically no significance at late time.)

Such "small" objects are ideally suited for testing in a facility such as AURORA. They can easily be transported to, and mounted in, the test cell. However, the limitations discussed earlier--the slow pulse rise and fast pulse decay--make certain coupling behavior impossible to observe using AURORA in its conventional, unmodified state. The authors have, however, greatly enhanced the usefulness of AURORA for such small-object testing by introducing into the test cell auxiliary sources of electromagnetic excitation.[18] For very small objects, this may simply take the form of a pair of thin aluminum plates to which a high-voltage pulse is applied (fig. 4). In this way, the proper electric field is reproduced in the enclosed volume, and the AURORA pulse serves the function of producing the radiation-induced conductivity and Compton current. The plates are made of aluminum foil in order to minimize shielding, space-charge effects, and other evidences of interaction between the radiation pulse and the auxiliary simulator. If the proper electric field is reproduced in this way, useful coupling measurements can be made on small systems--measurements which include the effects of time-varying air conductivity. (The conductivity pulse, though it cannot precisely reproduce the threat conductivity pulse, does (1) permit the experimenter to learn more about the "basic physics" of such interactions, and (2) provide a test bed for validation of theoretical techniques--both analytical and numerical--for predicting SREMP response of simple geometric forms, and, most importantly, of systems.)

In undertaking a simulation such as the one described above, one must take care to maintain control over the environment by minimizing any interaction between the conductivity pulse and the auxiliary simulator. In the parallel-plate structure described above, the interaction mechanism is simply that the radiation produces a conductive path across the plates--i.e., a time-varying conductance appears in parallel with the capacitance characterizing the simulator structure. This time-varying conductance has the value:

$$G(t) = \frac{C}{\varepsilon} \sigma(t) ,$$

where C is the plate capacitance. One can ensure that the appearance of this shunting conductive path produces only minimal distortion of the driving electric field seen between the plates, by incorporating into the structure a low-value resistor, R, such that:

$$R < \frac{\varepsilon}{C\sigma_{max}}$$

(See figure 4, where the 10-Ω resistor serves this purpose.)

The authors have used the approach described above, using two 8' x 4' aluminum foil plates separated by 2 ft, to excite a helical slow-wave structure intended to model a long cable or wire over ground. A 100-kV, 20-ns rise time, 2-μs fall-time pulser (built by Pulsar) was used to drive the system. The behavior of the simulator was monitored by a number of current sensors, by two single-ended E-field sensors mounted on the grounded plate, and by a balanced sensor suspended between the two plates.

Of course, the parallel-plate structure is in fact not a capacitor, but has some transmission-line characteristics. For the small simulator described above, these can reasonably be neglected, but for the testing of larger objects a larger structure is required. Work has been done in AURORA using a 10-m-long (3-m plate separation) structure,[17,18] whose inductive behavior cannot safely be ignored (fig. 5). The circuit of figure 6 is a more appropriate approximation for this case. An example of the radiation-induced distortion seen in such a structure is shown in figure 7.[17,18] The inductance

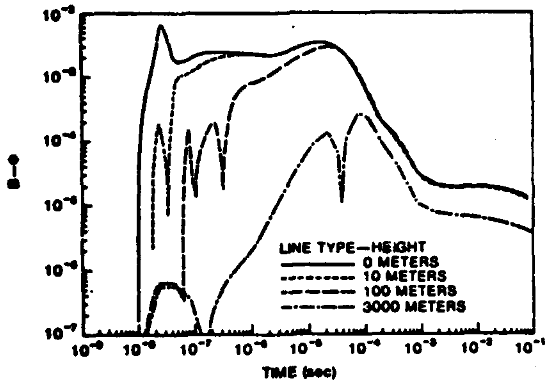


Figure 3. Typical time evolution of the magnetic field at various altitudes in the source region.

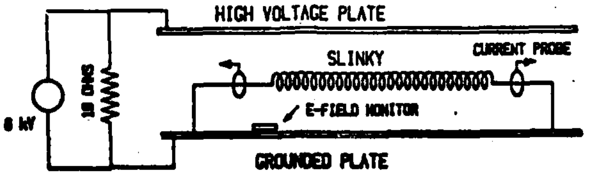

HIGH VOLTAGE PLATE
SLINKY
CURRENT PROBE
E-FIELD MONITOR
8 kV
10 OHMS
GROUNDED PLATE

Figure 4. Simple parallel plate capacitor auxiliary field producer. The auxiliary field producer compensates for the tendency of the test cell walls to short out the electric fields. The "slinky" is, in this case, the test object.

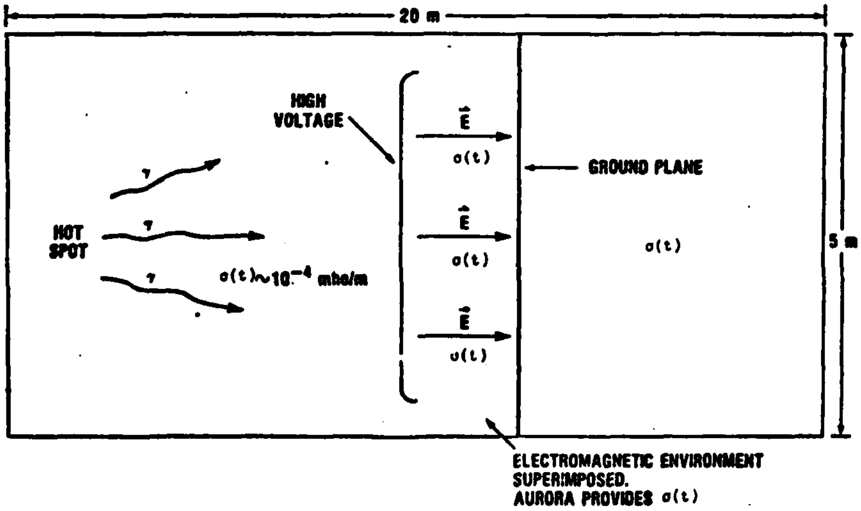

20 m
HIGH VOLTAGE
E
σ(t)
GROUND PLANE
HOT SPOT
γ
γ
γ
σ(t)~10⁻⁴ mho/m
E
σ(t)
σ(t)
E
σ(t)
5 m
ELECTROMAGNETIC ENVIRONMENT SUPERIMPOSED.
AURORA PROVIDES σ(t)

Figure 5. Schematic view of a transmission line designed to produce an auxiliary threat-relatable field in the AURORA test cell.

present in the system, in attempting to maintain current at a constant level, causes a voltage overshoot when the air conductivity drops abruptly. The authors have proposed, but not fully tested, a scheme for dealing with this more difficult problem. The low-impedance shunt must now be distibuted rather than lumped, and, of course, it must also be insensitive to radiation. This is achieved by introducing a low-impedance non-ionizing "slave" transmission line (patent pending[28]) in parallel with the "master" working line. This technique is described in greater detail elsewhere.[17,18,28]

Examples of simple coupling measurements showing short antenna response in the 10-m parallel-plate line, with and without radiation-induced time-varying air conductivity, are shown in figure 8.[17,20] The effect of conductivity is clearly seen here. It damps out the resonant displacement-current-driven response, and at the same time, superimposes a conduction-current-driven response. The authors have devised and experimentally verified an equivalent circuit technique which can be used to describe the interaction of antenna structures to an electromagnetic pulse in a medium with time-varying conductivity.[16,17,19,20,23,24,27]

The Strategic Case

The above approaches using auxiliary EMP sources are useful ways of imposing the desired electromagnetic environment while simultaneously superimposing a time-varying conductivity. (It should be noted that this isolation comes at a price--the "wastage" of a considerable fraction of pulser energy.) However, if the conductivity level gets too high (such as might be required for a deep source region or strategic simulation), the conductive air will begin to shield the plates from one another. The relaxation time, ε/σ , must be kept relatively long (greater than, say, 10 ns) for the duration of the AURORA pulse. But, for a "strategic" level of 0.1 mho/m at peak, the relaxation time can drop as low as 0.1 ns. Under such conditions, it is said that "local effects dominate" (fig. 9), and even if one can maintain the proper voltage separation across the plates, the intervening medium will short it out.

It is useful to note that the condition that "local effects dominate" is equivalent to the condition that Maxwell's wave equation degenerates into the diffusion equation, through the operator substitution:

$$\varepsilon \frac{\partial}{\partial t} \rightarrow \varepsilon \frac{\partial}{\partial t} + \sigma \simeq \sigma \quad .$$

Then $\nabla^2 E = \mu\sigma\dot{E}$.

This applies when:

$$\sigma > \varepsilon \frac{\partial}{\partial t} ;$$

i.e., when

$$\frac{\varepsilon}{\sigma} < T ,$$

where T is the shortest rise or fall time to be found in the driving signal. (Naturally, there is an intermediate case in which neither conductive nor displacement current can be neglected, but consideration of the two extremes

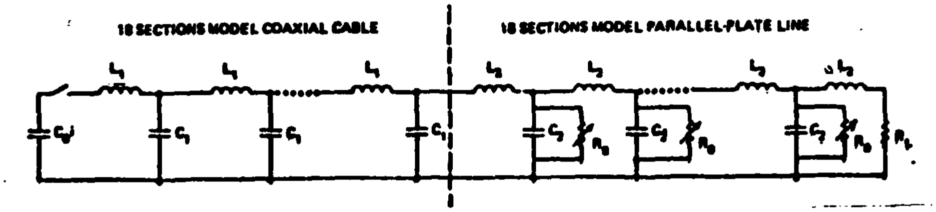

18 SECTIONS MODEL COAXIAL CABLE
18 SECTIONS MODEL PARALLEL-PLATE LINE
L_1
L_2
C_0
C_1
C_2
R_0
R_L

Figure 6. Schematic of transmission-line lumped parameter network model. The variable resistors represent the presence of air conductivity.

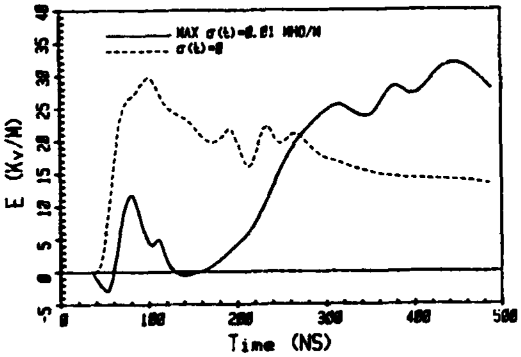


Figure 7. Extreme example of the effect of conductivity on the electric field in the Mark I SREMP simulator. First the air conductivity produces an undershoot in the line's electric field by shorting the transmission line; then the inductance of the line produces an electric field overshoot (inductive kick). The "slave line" of the proposed Mark II simulator is designed to reduce the undershoot and inductive kick.

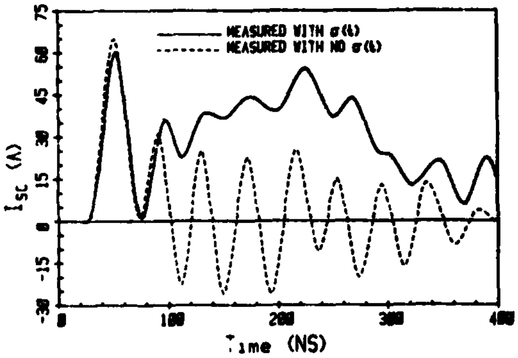


Figure 8a. Comparison of measured short-circuit monopole antenna response (2.42m) with and without time-varying air conductivity.

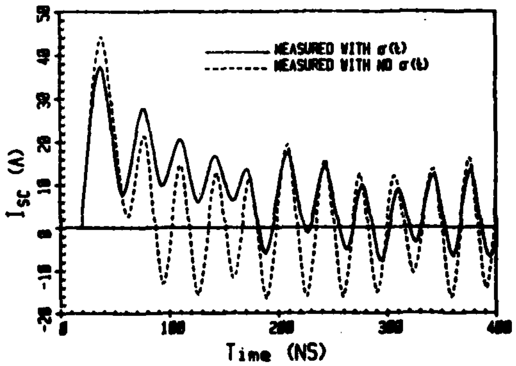


Figure 8b. Comparison of measured short-circuit monopole antenna response (2.21m) with and without time-varying air conductivity.

is useful for a qualitative understanding.)

When "local effects dominate," truly local drivers are needed to maintain a field. In the case of nuclear EMP, sustained Compton current performs this function. In the AURORA, the authors, in collaboration with S. Graybill, K. Kerris, D. Whittaker and other AURORA staff members, have achieved this effect by using direct electron injection.

AURORA's five megajoules of stored energy appear at the vacuum diode as a pulse of electron current. Ordinarily, a high-Z (usually tantalum) target is used to convert to x-ray energy. This is an inefficient process (assuming one's goal is to irradiate the test cell as heavily as possible) because:

(1) bremsstrahlung conversion efficiency at machine voltage is only about 6%, the remaining energy being lost as heat in the target, and

(2) the x-ray range is on the order of hundreds of meters, so that further energy is lost heating up the back wall.

If the target is removed and replaced with a 1/16th-in. steel vacuum-to-air interface, the electrons are released directly into the test cell and these inefficiencies are avoided. There is no conversion loss, and the electron ranges are of the order of the test cell length. Problems of self-consistent beam propagation (e.g., pinching and hosing) do not pose a threat, as has been determined in a series of AURORA experiments.[22]

In addition to thorough mapping of the electron irradiation of the test cell in a number of diode configurations and commbinations,[22] coupling measurements have been made (by the authors) on aluminum cylinders, both vertical and horizontal.[23] Examples of these measurements are shown in figure 10. Interpretation of the results is far from straightforward, since a number of drive mechanisms contribute. The most important of these are:

(1) displacement current (early time),
(2) conduction current,
(3) Compton current, and
(4) quasi-static space charge.

In addition to the higher irradiation of the room (as compared to bremsstrahlung-mode operation), the electron mode presents another significant benefit. The rise time appears to be faster. Evidence for this is shown in figure 11, where dose-rate measurements--taken with a Cerenkov detector in electron mode, and with a plastic scintillator in photon mode--are compared.

Conclusion

A good deal of thought and discussion is currently being generated on possible designs for new SREMP simulation techniques. While this is undoubtedly a healthy development, the authors would like to stress that the AURORA facility remains a significant and ever improving source of data

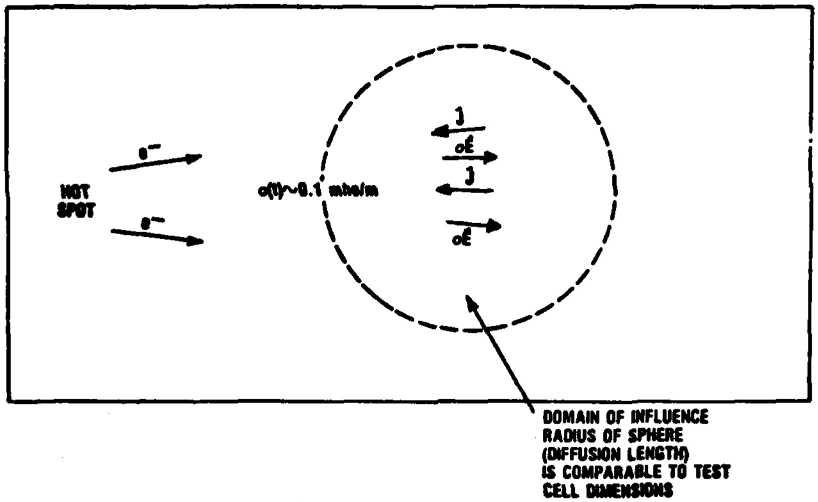

HOT SPOT
e⁻
e⁻
σ(t)~0.1 mho/m
J
σE
J
σE
DOMAIN OF INFLUENCE
RADIUS OF SPHERE
(DIFFUSION LENGTH)
IS COMPARABLE TO TEST
CELL DIMENSIONS

Figure 9. When electrons are injected into the AURORA test cell, the conductivity is so high that "local effects dominate".

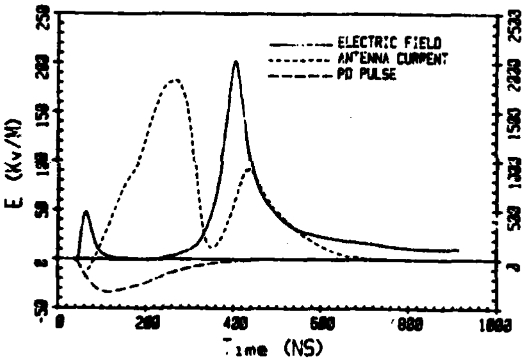


Figure 10. Cylindrical antenna current response, electric field measurement and photodiode measurement during direct electron injection AURORA shot.

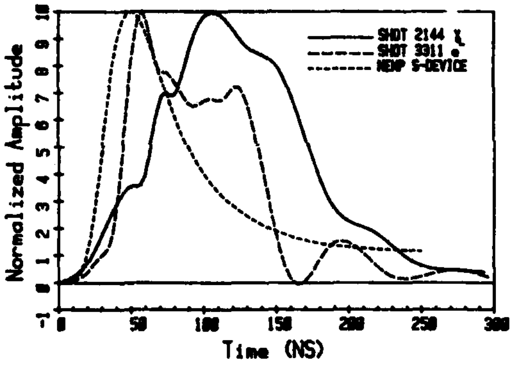


Figure 11. Rise time comparisons between various radiation sources. AURORA gamma mode pulse is from a single-source shot made during the July 1976 series. The electron mode pulse is from the March 1981 series.

relevant to SREMP environments and coupling. Tactical work using auxiliary sources and strategic work using the electron mode continue to provide a steady stream of information and new techniques which bear directly on problems of current interest.

Other existing radiation sources--such as HERMES II, which offers less energy, but which can be fired outdoors, eliminating the shorting effect of metal walls--are, and should be, under consideration for use in SREMP testing. Also, development of new radiation-source concepts is taking place at Harry Diamond Laboratories[31] and elsewhere. All these alternative sources can, and no doubt will, be used in conjunction with the techniques--auxiliary sources and electron mode--outlined in the above. However, for the time being, AURORA still dominates the SREMP scene.

REFERENCES

1. Stratton, J. A., Electromagnetic Theory, McGraw-Hill Book Company, New York, 1941.

2. C. L. Longmire, Close EMP Effects Lectures. LAMS-3072 and 3073, Los Alamos Scientific Laboratory, Los Alamos, NM, 1964. (Unpublished)

3. C. L. Longmire, Theory of the EMP from a Nuclear Surface Burst, LANC-R-8, Los Alamos Nuclear Corporation, Los Alamos, NM, 1970.

4. C. L. Longmire, On the Electromagnetic Pulse Produced by Nuclear Explosions, IEEE Trans. on Ant. and Prop., Vol., AP-26, No. 1, Jan. 1978, p. 3.

5. B. Bernstein and I. Smith, AURORA, An Electron Accelerator, IEEE Trans. on Nucl. Sci., Vol. NS-20, No. 3, June 1973.

6. Stewart Graybill, AURORA Facility Memo, Harry Diamond Laboratories, Adelphi, MD 20783.

7. Hans Fleischmann, High Current Electron Beams, Physics Today, May 1975, Vol. 28, p. 34.

8. Gerold Yonas, Fusion Power with Particle Beams, Scientific American, Nov. 1978, p 50.

9. J. F. W. Dietz, G. Merkel, and D. Spohn, Radiation-Induced Coupling to a Truncated Cylinder within a Cylinder, IEEE Trans. Nucl. Sci., Vol. NS-23, Dec. 1976, pp. 1982-1985.

10. J. N. Bombardt, Jr., J. F. W. Dietz, G. Merkel, and D. Spohn, Status of the Tactical Environment Multiple Systems Evaluation Program (TEMSEP), Harry Diamond Laboratories Report HDL-PR-77-2, Oct. 1977.

11. G. Merkel and D. J. Spohn, C. L. Longmire, and W. F. Crevier, An

Bushell, Manriquez, Merkel, & Scharf

Equivalent Circuit Analysis of the HDL Concentric Cylinder Experiments in AURORA, IEEE Trans. Nucl. Sci., Vol. NS-24, No. 6, Dec. 1977, pp. 2411-2415.

12. W. F. Crevier, E. D. Kalasky, Evaluation of AURORA as a Tactical Source Region Simulation, AURORA EMP Memo 21 (internal Mission Research Corporation Memorandum, Mission Research Corporation, Santa Barbara CA), June 1980.

13. W. F. Crevier, C. L. Longmire, G. Merkel, and D. J. Spohn, Air Chemistry and Boundary Layer Studies with AURORA, IEEE Trans. Nucl. Sci., Vol. NS-24, No. 6, Dec. 1977, pp. 2406-2410.

14. T. A. Tumolillo, J. P. Wondra, J. N. Bombardt, G. Merkel, and D. Spohn, PRES-D: A Computer Code for the Self-Consistent Solution of the Maxwell-Lorentz Three-Species Air-Chemistry Equations in Three Dimensions, IEEE Trans. Nucl. Sci., Vol. NS-24, No. 6, Dec. 1977, pp. 2456-2460.

15. R. P. Manriquez, G. Merkel, and D. Spohn, Modification of the AURORA Electromagnetic Environment: Experiment and Interpretation, Harry Diamond Laboratories Report HDL-PR-79-5, Aug. 1979.

16. R. P. Manriquez, G. Merkel, W. D. Scharf, and D. Spohn, Electrically-Short Monopole Antenna Response in an Ionized Air Environment--Determination of Ionized Air Conductivity, IEEE Trans. Nucl. Sci., Vol. NS-24, No. 6, Dec. 1979, pp. 5012-5018.

17. G. Merkel, W. D. Scharf, and D. Spohn, Use of the AURORA X-Ray Machine as a Source-Region Simulator and Antenna Coupling Analysis Facility, Proceedings of NATO Defense Research Group Seminar, Shrivenham, UK, 15-17 Apr. 1980.

18. M. Bushell, R. Manriquez, G. Merkel, W. Scharf, and D. Spohn, Source-Region EMP Simulator--A Parallel-Plate Transmission Line in the AURORA Test Cell, IEEE Trans. Nucl. Sci., Vol. NS-27, Dec. 1980, pp. 1834-1838.

19. M. Bushell, R. Manriquez, G. Merkel, W. D. Scharf, and D. Spohn, Capacitively and Inductively Loaded Antennas in a Source-Region EMP Environment with Time-Varying Conductivity, IEEE Trans. Nucl. Sci., Vol. NS-27, Dec. 1980, pp. 1857-1863.

20. M. Bushell, R. Manriquez, G. Merkel, W. D. Scharf, and D. Spohn, An Equivalent Circuit for a Linear Oscillating Antenna Immersed in an Environment with Time-Varying Conductivity, IEEE Trans. Nucl. Sci., Vol. NS-27, Dec. 1980, pp. 1857-1863.

21. G. Merkel, W. D. Scharf, and D. Spohn, Source-Region EMP Simulator--A Parallel-Plate Transmission Line in the AURORA Test Cell, Proceedings of DNA Symposium on EMP Simulation and System Hardening, Monterey, CA, Oct. 1980.

22. M. Bushell, S. Graybill, K. Kerris, G. Merkel, W. D. Scharf, and D. Whittaker, The Direct Injection of Electron Pulses into Air--A SREMP Simulation Tool, IEEE Trans. Third Int. Pulsed Power Conference, June 1981.

23. M. Bushell, R. Manriquez, G. Merkel, and W. D. Scharf, Simulated Deep Source-Region EMP Coupling to Cylindrical Objects, IEEE Trans. Nucl. Sci., Vol. NS-28, No. 6, Dec. 1981, p. 4457.

24. M. Bushell, R. Manriquez, G. Merkel, W. D. Scharf, Measurement and Calculation of the Short-Circuit SREMP Response of Vertical Oscillating Helical Antennas and Oscillating Loop Antennas, IEEE Trans. Nucl. Sci., Vol. NS-28, No. 6, Dec. 1981, p. 4495.

25. L. Ambrose, M. Bushell, C. Kenyon, G. Merkel, W. D. Scharf, Microwave Measurements of Conductivity in Air and Other Gases, IEEE Trans. Nucl. Sci., Vol. NS-28, No. 6, Dec. 1981, p. 4426.

26. M. Bushell, R. Gray, C. Kenyon, R. Manriquez, G. Merkel, and W. D. Scharf, Strategic and Tactical Source-Region EMP Simulation Techniques, DNA Conf. on Instrumentation for Nucl. Weapons Effects, March, 1982.

27. G. Merkel, W. D. Scharf, and D. Spohn, Simulation of Large Coupling Structures in a Source-Region EMP Environment, DNA Symposium on EMP Simulation and System Hardening, Monterey, CA, Oct. 1980.

28. H. J. Longley and C. L. Longmire, Development and Testing of LEMP1, Los Alamos Scientific Laboratory LA-4346, April 1970.

29. H. A. Sandmeir, S. A. Dupree, C. E. Hansen, Electromagnetic Pulse and Time-Dependent Escape of Neutrons and Gamma Rays from a Nuclear Explosion, Nucl. Sci. and Eng., Vol. 48, 1972, pp. 343-352.

30. G. Merkel and W. D. Scharf, "Source-Region Electromagnetic Pulse Simulator", U. S. Patent Application #252737.

31. A. G. Stewart, "CAMELOT--A Novel Concept for a Multiterrawatt Pulse Power Generator for Single Pulse, Burst or Repetition Rate Operation", Harry Diamond Laboratories Report HDL-SR-81-1, April 1981.

CONTROL OF PARASITIC CURRENTS ON RADIATING SYSTEMS

DONN V. CAMPBELL, PH.D.
US ARMY COMMUNICATIONS-ELECTRONICS COMMAND
FORT MONMOUTH, NEW JERSEY 07703

I. INTRODUCTION:

Military communication antennas operating in the 2-30 MHz high frequency (HF) band, 30-90 MHz very high frequency (VHF) band, and the 200-400 MHz ultra high frequency (UHF) band are installed on vehicular, shipborne, airborne, manpack, or fixed platforms. Short, medium and long range radio is supported by HF, Net Radio is serviced by VHF, and ground-to-air communication employs UHF. The antenna engineer is confronted with the task of designing small efficient radiating systems having wide bandwidth and predictable impedance and radiation characteristics. In the past, wide bandwidth has been achieved at HF and VHF by band switching and broadband matching circuits. However, new requirements for wide instantaneous bandwidth (no tuning) antennas have been imposed with the advent of frequency hopping radios.

Antenna interaction (mutual coupling) and electromagnetic compatibility (EMC) are major concerns when many antennas and communication systems are collocated in a confined area, as in a command post or an air traffic control facility. System performance may be unpredictable due to antenna mutual coupling effects. The purpose of this paper is to demonstrate, by computer analysis, and validating experimental measurements, that extraneous parasitic radio frequency (rf) currents on radiating systems can impair performance, causing field pattern distortion, crosstalk interference, and unpredictable behavior with changes in operating frequency. It is shown that parasitic currents can be suppressed in both narrowband and wideband radiating systems. The techniques discussed are practical and are applicable to field use where accepted procedures for antenna installation are sometimes set aside because of lack of understanding, or for expediency.

This paper is organized as follows: the horizontal dipole above ground is studied, and it is shown that parasitic currents on the sheath of the coaxial feed line can be suppressed. Similarly, the inverted-V dipole is shown to have significant parasitic feed line currents which can also be eliminated. Extraneous rf currents on the mast supporting a vertical dipole are shown to be substantially reduced over a narrow frequency range by the addition of a quarter-wave detuning stub and over a much wider frequency range by means of a broadband cable choke. Associated radiation patterns confirm the improvement obtained when parasitic currents are suppressed. A broadband colinear dual antenna providing more than 35 decibels (dB) of inter-dipole isolation is described. The high isolation results from the colinear antenna arrangement, and by virtue of the special feed line "Isolator Section" which acts as a band elimination filter for the parasitic currents induced on the feed line of the upper dipole. This antenna system allows duplex operation free from crosstalk interference in an air traffic control facility.

II. PARASITIC CURRENTS ON HORIZONTAL DIPOLES

The horizontal dipole center fed by coaxial line is extensively used. It is known that unequal electromagnetic forces on each half of this antenna cause a condition of imbalance which may result in unequal currents in the antenna arms. The feed line coupled to the antenna may support parasitic currents of appreciable magnitude. In an earlier experimental study, the currents on a horizontal dipole above ground were measured (1). This earlier study demonstrated conclusively, that significant current imbalance may occur. It also showed that current balance in the antenna arms could be improved by simply connecting a cable choke between the dipole and the feed line. In essence, the cable choke acts as an effective high impedance circuit which interrupts the rf current flowing on the outer surface of the feed line.

For comparison purposes, the experimental horizontal dipole was modeled on the Numerical Electromagnetic Code (NEC), a computer program for analyzing the electromagnetic response of antennas (2). The antenna is shown in Fig. 1. To facilitate measurements of current distribution, the experimental dipole was installed seven feet above the ground, and it was modeled on the computer accordingly. In the computer model perfectly conducting ground was assumed and the bottom of the feed line was assumed to be connected to the ground.

Although the length of the modeled antenna is a half wavelength at 4 MHz, such an antenna might be operated in the field at other frequencies without bothering to adjust its length to resonance. To investigate this possibility, the current amplitude distribution at 2 MHz was determined, and is shown in Fig. 1 by the dashed line curve. The experimental and computed distributions are similar. It is seen that the current imbalance is pronounced and that the feed line supports a parasitic current of appreciable amplitude on its outer surface.

As stated earlier, a cable choke was used to suppress parasitic current on the feed line and to improve current balance on the arms of the experimental dipole antenna. A cable choke consists of coaxial cable shaped into a helical coil, as shown in Fig. 2(a), or in the form of a toroid. In the experimental study, the cable choke consisted of thin coaxial cable wound on several high permeability ferrite toroids. Regardless of the construction, however, the cable choke impedance between points A and B at the outer surface of the transmission line (see Fig. 2(a)) is essentially equivalent to a high impedance circuit consisting of an inductance L connected in parallel with capacitance C, as shown in Fig. 2(b). When losses are neglected, the reactance of the choke is, to good approximation, given by $X = 2\pi fL / (1 - (f/fo)^2)$ where the self resonance frequency is $fo = 1/2\pi \sqrt{LC}$. In practice, cable choke losses will be present and should be minimized.

In wide bandwidth applications, it is practical to design the cable choke so that it resonates at the geometric mean frequency $f_o = \sqrt{f_L f_U}$ where f_L and f_U denote, respectively, the lower and upper frequency limits. When so designed, the reactances at the frequency limits will be equal in magnitude. The bandwidth of the cable choke is inversely proportional to the self capacitance. Thus, to maximize the isolation bandwidth, the cable choke should be designed for maximum inductance and minimum capacitance.

When the cable choke was connected between the dipole and the feed line, in the manner shown in Fig. 3, it was found that the parasitic current on the feed line was essentially eliminated, and the antenna current balance was improved. The impedance of the cable choke used with the experimental diple antenna was (3120 + j4550) ohms at 2 MHz (1). The effect of this choke was accounted for in the NEC computer model by inserting this impedance between one arm of the dipole and the feed line. It is seen that the calculated and measured currents are comparable. Although not shown here, similar improvements in the current distribution are effected at 4, 6, and 8 MHz, demonstrating that the cable choke provides an effective means of suppressing parasitic currents over a broad frequency range.

III. PARASITIC CURRENTS ON INVERTED-V DIPOLES

The inverted-V dipole with drooping arms is an easily erected antenna requiring only one support at its midpoint. As in the previous example, the NEC computer program was used to obtain the current distributions with the feed line connected directly to the antenna and with a cable choke interposed. The inverted-V shown in Fig. 4 is resonant at approximately 7 MHz. At resonance, currents in the antenna arms are well balanced and the feed line excitation is minimal. However, if the antenna is operated

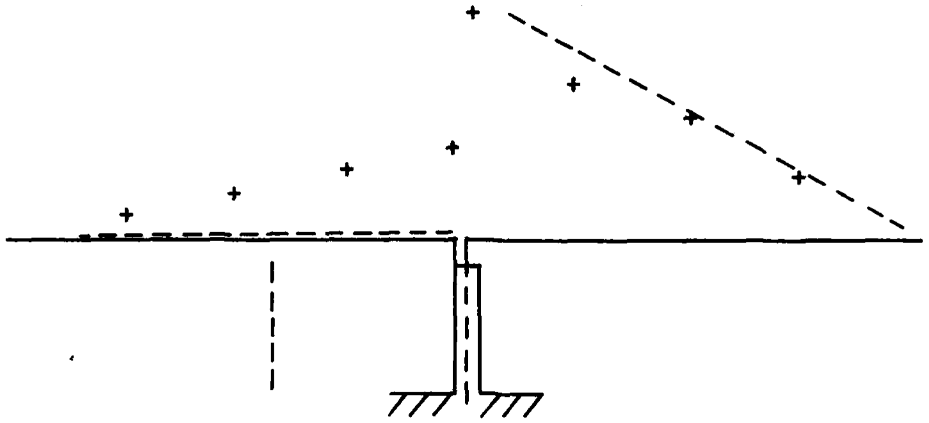
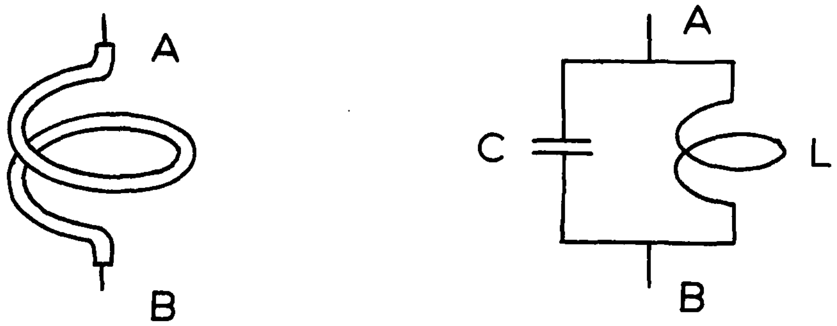


Fig. 1. Horizontal dipole with unbalanced currents;
--- computed + measured



(a) (b)

Fig. 2. (a) Cablechoke (b) Equivalent circuit

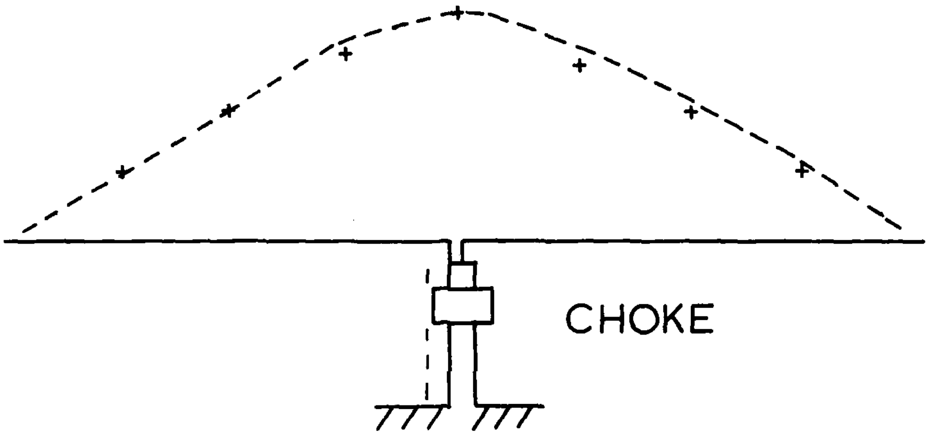


Fig. 3. Horizontal dipole with balanced currents;
--- computed + measured

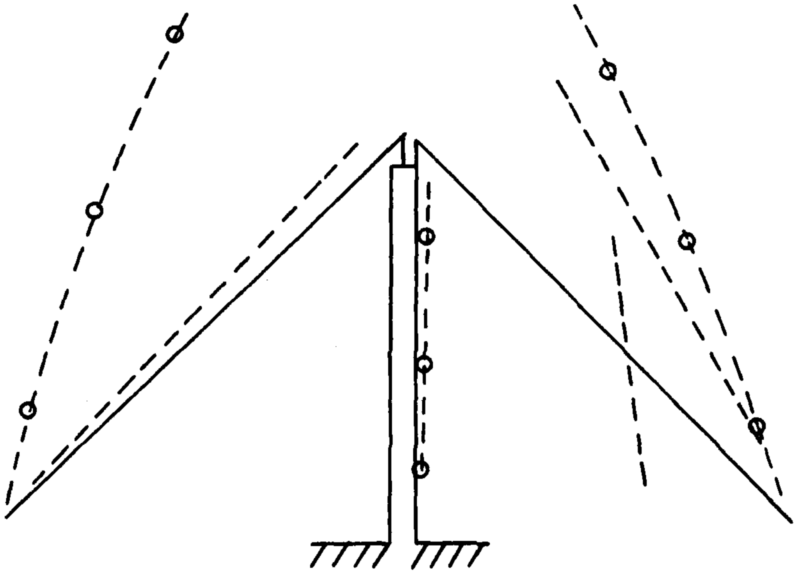


Fig. 4. Currents on inverted-V dipole
--- without cable choke and -o- with cable choke

at frequencies above or below resonance, current imbalance and feed line excitation will increase. The current amplitude distribution obtained at 4 MHz, for example, confirms this expectation. See Fig. 4. Note that in this computer model, the lower end of the feed line is assumed to be connected to perfectly conducting ground.

The inverted-V dipole with a cable choke connected in series with the feed line was also investigated. The impedance of the cable choke at 4 MHz is assumed to be (3935 + j2937) ohms. The lower end of the feed line in the model does not contact the ground, simulating an ungrounded installation. In the computer model "average" ground having a relative permittivity of 15 and conductivity of .005 mhos per meter was assumed. It is seen in Fig. 4 that the cable choke eliminates parasitic current on the sheath of the feed line and balances current in the antenna arms. It is interesting to note that the inverted-V feed-point impedance is (30.1 - j857) ohms with the cable choke present and is (542 + j1942) ohms without the cable choke. This change in impedance is attributed to the markedly different current distributions obtained for the two cases. If a metal support mast is used, an insulator should be provided at its top to prevent "short-circuiting" the cable choke.

IV. PARASITIC CURRENTS ON MAST-MOUNTED DIPOLES

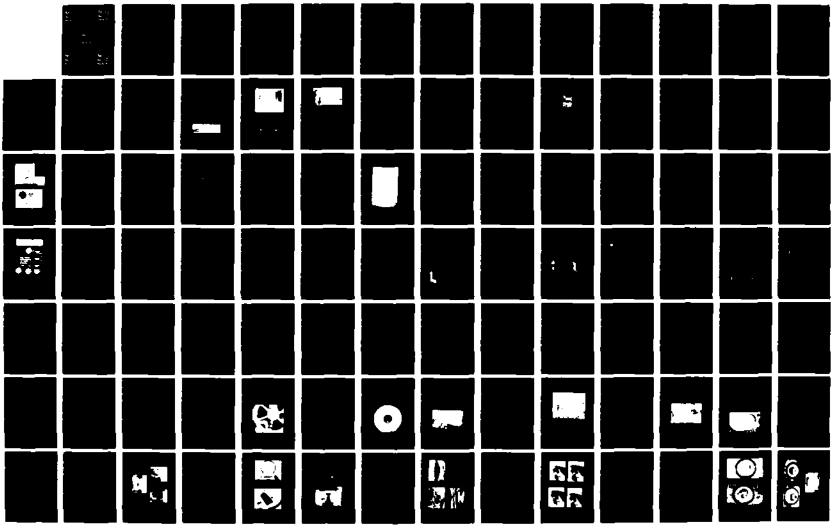
Mast-mounted vertical dipoles are extensively used at VHF and at UHF. The dipole can be electrically decoupled from its support by attaching a quarter-wavelength detuning stub to the mast, or to the feed line as shown in Fig. 5(a). Provided that the stub is properly dimensioned, it introduces a high impedance at the lower end of the dipole cutting off the line current or mast current below. The current amplitude on the resonant stub is maximum at the shorted end. The section of feed line or mast adjacent to the stub carries both the antenna current and the transmission line current. The transmission line currents are in opposite directions in the stub and adjacent mast so that the resultant current is small. Other applications of stubs appear in the literature (3, 4).

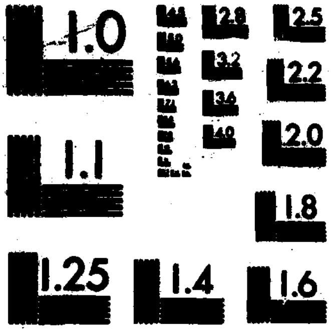
If the length of the stub is somewhat different from a quarter-wavelength at the operating frequency, it is then ineffective because its impedance is too low to enforce a current minimum on the mast. This is illustrated in Fig. 5(b), where the non-resonant stub has little effect. Evidently, detuning stubs are effective but in narrowband applications.

When the mast currents are suppressed, as in Fig. 5(a), it is found that the field strength on the horizon increases. Conversely, the field strength on the horizon is reduced when parasitic current is not suppressed. The vertical radiation patterns for these two cases are shown in Fig. 6. For the assumed conditions, the power gain on the horizon is increased by more than 6 dB by suppressing the mast currents. This, of course, trans-

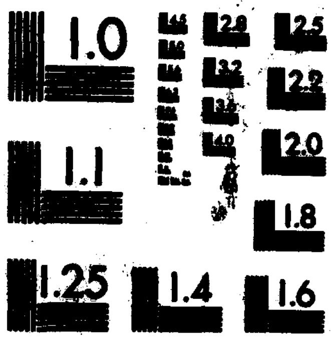
AD-A120 811 PROCEEDINGS OF THE 1982 ARMY SCIENCE CONFERENCE HELD AT THE UNITED STATES..(U) DEPUTY CHIEF OF STAFF FOR RESEARCH DEVELOPMENT AND ACQUISITIO.. 18 JUN 82 3/6

UNCLASSIFIED F/G 5/2 NL

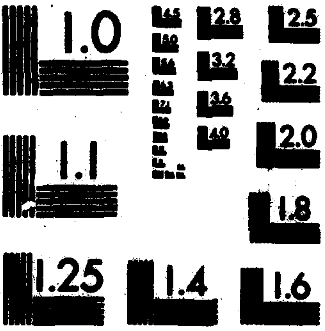



1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

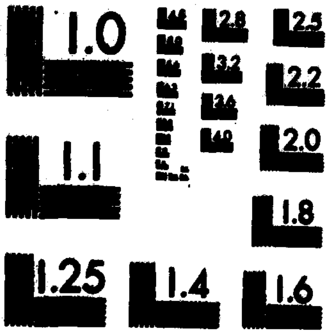
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

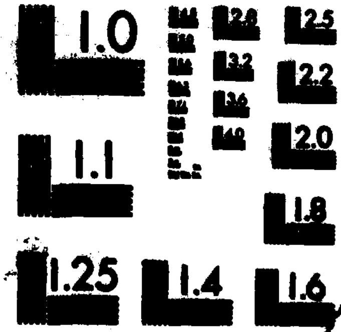
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

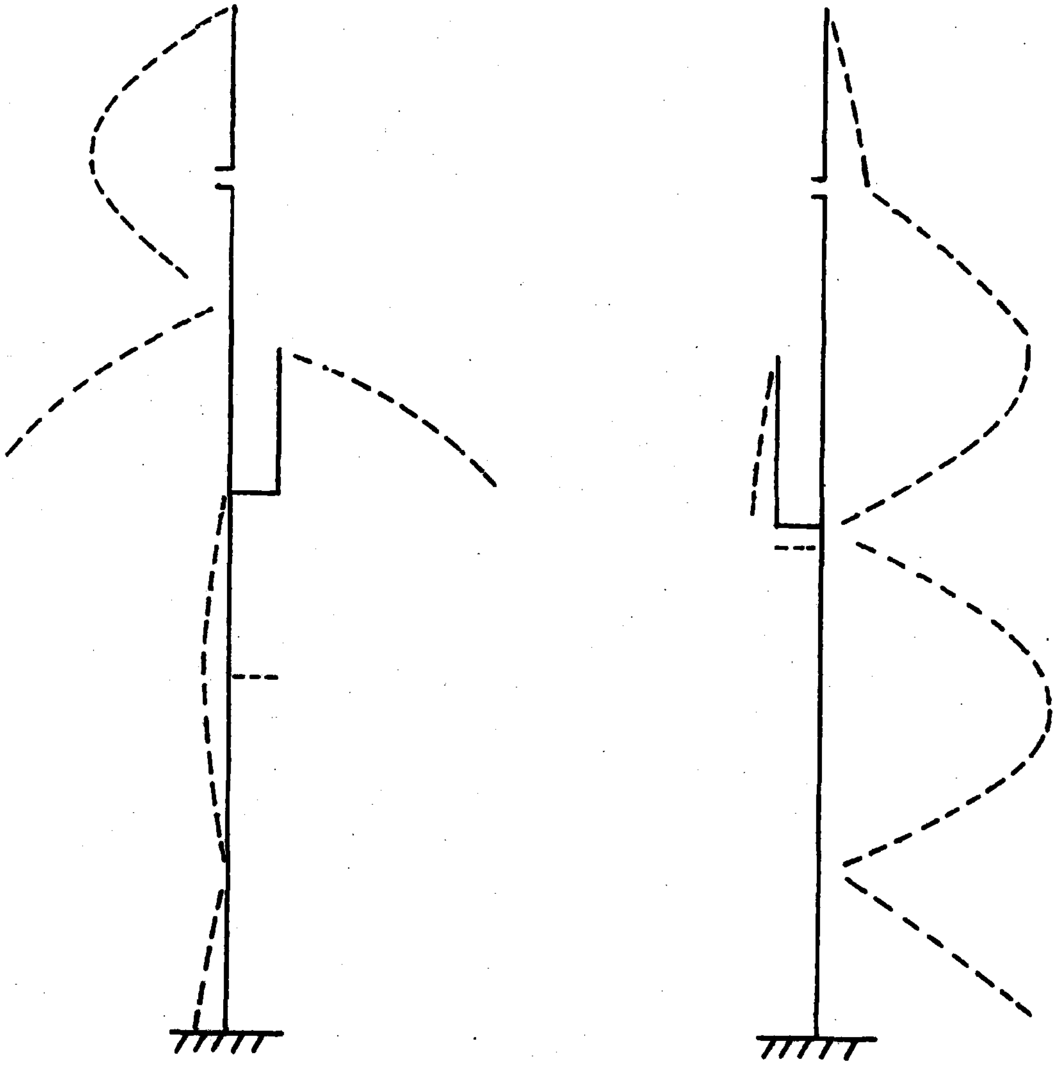


Fig. 5. Currents on Mast-mounted dipole
(a) with resonant stub (b) with non-resonant stub

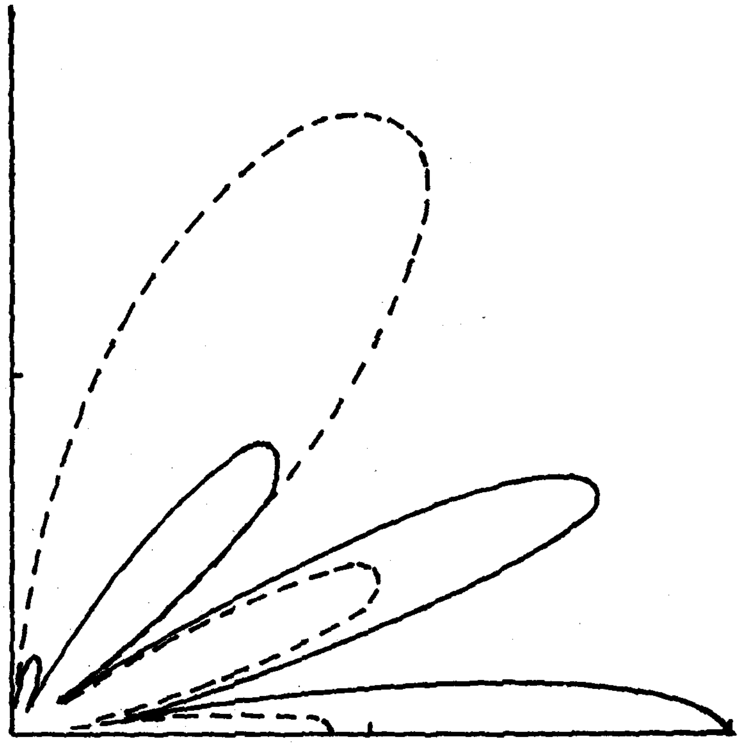


Fig. 6. Vertical radiation patterns of mast-mounted dipoles;
—— non-resonant stub —— resonant stub

lates into greater operating range.

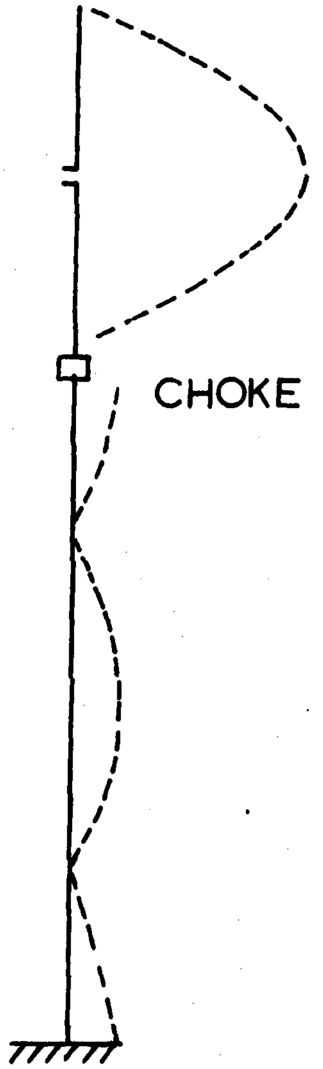
The dipole can also be decoupled from the mast and from the feed line by a cable choke as shown in Fig. 7. To test this hypothesis, a UHF dipole was modeled, using NEC, and the current distribution at 300 MHz determined. The cable choke was assumed to be equivalent to a parallel LC circuit (see Figs. 2(a) and 2(b)) with constants appropriate to UHF. The current is essentially confined to the half wavelength dipole and is relatively weak on the mast or feed line. Even better results could be obtained if the cable choke were resonant at the operating frequency. In this example the LC constants chosen cause the choke to resonate at 258 MHz, which is somewhat lower than the operating frequency. The radiation pattern would be similar to that already shown (Fig. 6) for the antenna with resonant detuning stub.

V. PARASITIC CURRENTS ON A COLINEAR DUAL ANTENNA SYSTEM

In duplex communication systems, the mutual coupling of the collocated transmitting and receiving antennas can result in crosstalk interference. It is well known that radiation coupling between two dipole antennas is minimized when they are mounted colinearly, and sufficiently spaced. A second coupling effect, due to parasitic currents induced on the feed line of the upper dipole, is more difficult to control than radiation coupling. With radiation coupling minimized, this second coupling effect predominates.

It has been demonstrated recently that these parasitic currents can be suppressed over an octave frequency range by means of an "isolator section" incorporated in the feed line of the upper antenna (5). The feed line isolator section consists of a sequence of high impedance cable chokes placed in the coaxial line at quarter-wavelength intervals. It behaves electrically as a band elimination filter which reduces the interantenna coupling caused by currents on the transmission line sheath.

The colinear dual antenna is shown in Fig. 8. The isolator section in this antenna has five cable chokes. The antennas are independently fed dipoles. A computer model of this antenna yields the current distribution shown when the upper dipole is excited, and the lower dipole is terminated in 50 ohms. It is seen that current is essentially restricted [illegible] driven dipole, and is very weak along the rest of the antenna system. [illegible] isolation between the upper and lower dipole is defined as the frac[illegible] the total power radiated by one dipole which is intercepted by the other dipole. The interdipole isolation of the antenna system (Fig. 8) was found to be at least 35 dB from 200 to 400 MHz, indicating that radiation coupling is the limiting factor for the given antenna spacing.



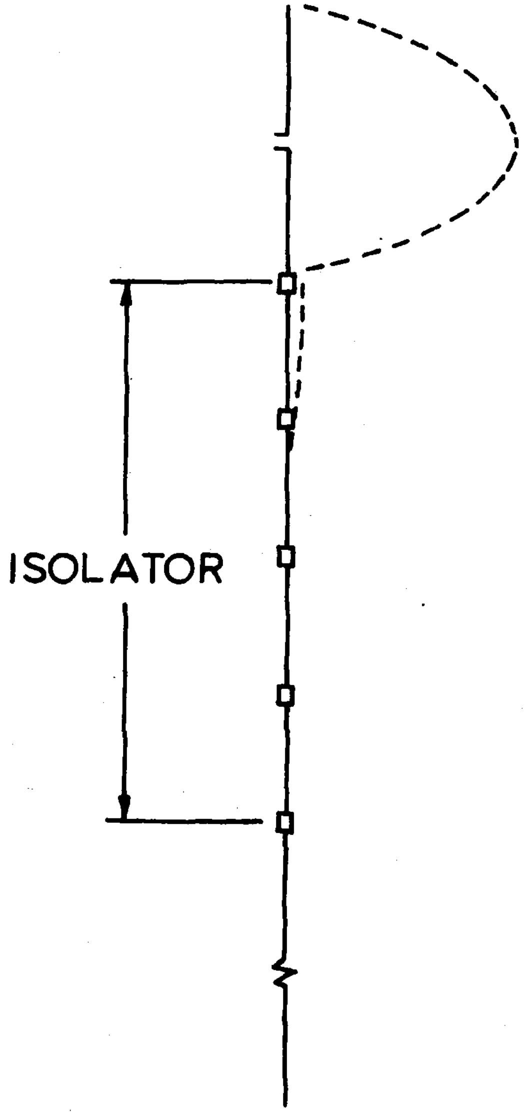


Fig. 7. Mast-mounted dipole with cable choke.

Fig. 8. Colinear dual antenna with isolator section.

VI. CONCLUSIONS

In this paper, it has been clearly shown that induced parasitic rf currents can seriously degrade the performance of a radiating system; yet the significance of these undesired spurious currents has been largely ignored by antenna designers, and field users alike. It has been shown that parasitic currents can be effectively suppressed in both narrowband and wideband antenna systems by means of quarter-wavelength stubs (resonance method) and broadband cable chokes. Although effective current suppression techniques exist, further study is necessary to establish engineering design criteria and to extend state-of-the-art current suppression techniques to multi-octave antenna systems.

REFERENCES:

(1) J. P. Muro, "Investigation of Unbalanced and Force Balanced Dipole Antennas," Master's Project, Newark College of Engineering, Newark, NJ, 1968.

(2) J. G. Burke, A. J. Poggio, J. C. Logan, and J. W. Rockway, "NEC-Numerical Electromagnetics Code for Antennas and Scattering," in IEEE 1979 Int. Symp. on Antennas and Propagat., Vol. 1, pp. 147-150, Seattle, WA.

(3) R. W. P. King, H. R. Mimno, and A. H. Wing, TRANSMISSION LINES, ANTENNAS AND WAVE GUIDES, New York, Dover Publications, Inc., 1965.

(4) J. S. Belrose, "On Minimizing the Effects of Re-Radiation from Power Transmission Lines on the Radiation Pattern of MF-AM Broadcasting Antenna Arrays," IEE Second Int. Conf. on Antennas and Propagat., Part 1, pp. 33-37, York, U.K., 1981.

(5) D. V. Campbell, S. Fich, and F. K. Schwering, "Suppression of Parasitic Currents on Feed Lines of Colinear Dual Antenna Systems," IEEE Trans. Antennas Propagat., Vol. AP-28, pp. 658-662, Sept. 1980.

DESIGN OF MICROSTRIP LINEAR ARRAY ANTENNAS BY COMPUTER

MORRIS CAMPI, PH.D.
U.S. ARMY ELECTRONICS RESEARCH AND DEVELOPMENT COMMAND
HARRY DIAMOND LABORATORIES
ADELPHI, MARYLAND 20783

Introduction

Research in microstrip antenna technology and design is currently being developed to meet the present and future needs of military radar and fuze system requirements. Configured in either single- or multiple-lattice structures, microstrip has the advantages of ease in fabrication, low cost, light weight, and structural conformability. When designed to have either a series or a corporate feed network, the lattices become antenna arrays that can be used for transmitting and receiving signals from microwave through millimeter frequencies into the lower regions of the near-millimeter wave portion of the spectrum. However, the present state of the art is at the same stage of development as existed for waveguide slot array designs in the late 1950's and early 1960's. This stage is a "cut-and-try" period in which an empirical design takes from weeks to months of fabrication and testing before satisfactory results are obtained; even then, one can never be sure of achieving the optimum design to meet the system's specific requirements.

Today, computer aided design codes can be developed to help antenna engineers in their design of linear or planar array microstrip antennas. This report describes one such code that was developed by using both empirical data and a theoretical model. One major feature of this code is that it allows the designer to vary any one or more of the antenna design parameters and to observe the results, such as changes in the radiation pattern or amplitude distribution displayed at the computer terminal screen. This application provides the tool required by the engineer to determine in a matter of seconds the prime factors affecting his design and to make parametric adjustments or sensitivity studies to help understand and develop his final product. Once a design is established, the engineer is presented with a set of coordinate numbers that can be placed on magnetic or paper tape, which can then be used to fabricate the antenna on printed circuit board without the need for drawings or art-layout work and thus simplify the engineering process and reduce the time spent.

The computer program reported will continuously evolve. As better data or refinements in the analytical models are produced, they will be incorporated into the program.

Microstrip Linear Array Antenna

Microstrip arrays, like other array designs, have a power feed arrangement that is classified as being either series or corporate feed. A series-feed array is designed so that each element connects with the next element in series with the input power or receiver being applied at one end of the array. In a corporate-feed array, each element is independent of each other, and a prescribed level of power is individually transferred to or received from each element by separate means or by use of a power divider network. The radiation characteristics of either type of antenna are essentially those produced by the electromagnetic fields that illuminate the antenna aperture. (The far-field radiation pattern is related to the Fourier transform of the aperture illumination.) A constant amplitude of illumination at each array element produces a $\sin^2 x/x^2$-like pattern in the far field with 13.2-dB side lobes below the main beam. Other distributions produce far-field patterns with other levels of gains, beam widths, and side lobes, each dictated by the antenna aperture distributions. Over the years, some specific distributions have been introduced and used by engineers to satisfy their systems requirements. These distributions have since become quite common in the industry.

Control of the aperture distribution in a corporate-feed array is fairly simple. However, the feed or power dividing network itself can become complicated and, if designed in microstrip, radiates and causes a distortion in the radiated power distribution. A series-feed array, however, has no independent feed network, but is power fed from one end. As the power traverses along the array, some power radiates from the first element and the rest continues until all the elements radiate the remaining power. The aperture distribution in this case is determined by the radiation characteristics of each element and their relation to all of the other element characteristics. A computer analysis of this type of array can be obtained once the single-element characteristics along with the array neighbor-to-neighbor interactions are understood.

Single Element Patch Model

The geometry of the single element microstrip patch is shown in figure 1. It consists of a thin rectangular conducting plate positioned over a conducting ground plane separated by a thin layer of dielectric material, where the material thickness is much less than one wavelength. By use of inexpensive printed circuit technology, the patch element is readily fabricated by using etching techniques applied to metal clad dielectric, usually Teflon or Teflon impregnated with fiberglass sandwiched between thin copper plating.

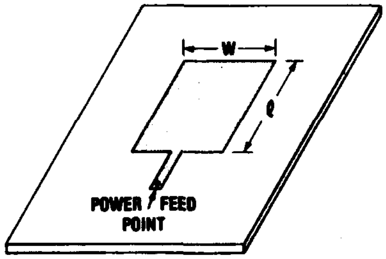


Fig. 1. Geometry of microstip patch element.

The patch width, W, determines the electrical admittance of the element, and the dimension ℓ (nearly one-half wavelength) determines the frequency at which resonance occurs. The element consists basically of two radiating slots perpendicular to the feed line and separated by a transmission line of very low impedance (1). At resonance, the element is an efficient radiator of electromagnetic energy. The resonant line length of the patch is slightly less than one-half wavelength because of phase modifications due to substrate thickness, fringing field capacitance, and the patch aspect ratio. It is calculated (2) as

$$\ell = c/\left(2f_o\sqrt{\varepsilon_e}\right) - 2\delta\ell \quad , \tag{1}$$

where c = 30, f_o is the resonant frequency in gigahertz,

$$\varepsilon_e = \frac{\varepsilon_d + 1}{2} + \frac{\varepsilon_d - 1}{2}\left(1 + \frac{12h}{W}\right)^{-1/2} \quad , \tag{2}$$

$$\delta\ell = 0.412h\,\frac{\left(\varepsilon_e + 0.3\right)\left(W/h + 0.264\right)}{\left(\varepsilon_e - 0.258\right)\left(W/h + 0.8\right)} \quad , \tag{3}$$

ε_e is the effective dielectric constant, ε_d is the dielectric constant, and h is the dielectric thickness.

The fields radiating from these slots have components parallel to the ground plane, which add in phase to give a maximum radiated field normal to the element. Figure 2 illustrates the geometry of a radiating slot.

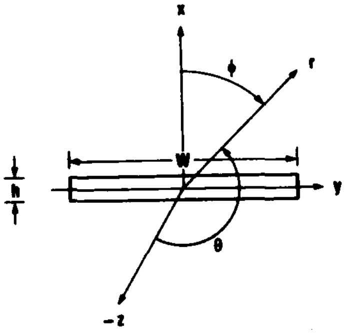


Fig. 2. Geometry of radiating slot.

The total power radiated from each slot is obtained by integrating the real part of the Poynting vector from a single slot over a hemisphere of large radius. It is derived as

$$P = \sqrt{\frac{\varepsilon}{\mu}}\frac{V_0{}^2}{\pi}\int_o^{\pi} \frac{\sin^2\left(\frac{\pi W}{\lambda_0}\sin\phi\right)}{\sin^2\phi}\cos^3\phi\, d\phi \quad , \tag{4}$$

where V_0 is the voltage across the slot of height h, which is very small, that is, $k_0h << 1$, where k is the propagation constant equal to $2\pi/\lambda_0$.

Microstrip Element Conductance

Since rigorous transmission line or two-port parameters are not available for a resonant, radiating microstrip patch, it is modelled as a lumped conductance and a phase delay. The delay is taken into account for the phasing of series-feed arrays; the combined power radiated by the two slots of a patch is the same as that dissipated by the conductance, G, having across it the voltage, V_0, at the center of the slot; hence, $G = P/V_0^2$ or $G = I/\pi\sqrt{\varepsilon/\mu}$, where

$$I = \int_0^{\pi} \frac{\sin^2\left(\frac{\pi W}{\lambda_0}\sin\phi\right)}{\sin^2\phi}\cos^3\phi\, d\phi \quad . \tag{5}$$

The integral, I, is solved by use of computer integration techniques. In the range $0.033 \leq W/\lambda_0 \leq 0.254$, conductance values have been measured and follow the relation (3)

$$G = 0.0162\left(\frac{W}{\lambda_0}\right)^{1.757} \quad . \tag{6}$$

Insertion Phase

Transmission line discontinuities produce phase shifts or delays in the propagation of energy down the transmission line. The impedance mismatch and line-to-patch width aspect ratio at the entrance and exit ports of each microstip patch produce phase delays to the propagating signal. Most importantly, mode structure and wave number along the longitudinal axis of the microstrip cavity can introduce a large insertion phase shift between input and output ports in addition to the desired phase shift of 180°. This effect does not alter the resonance condition of the patch, but increases the effective phase delay to the following patch in the array and thus changes the composite radiation pattern. For example, if all the patch widths were equal, the resulting constant insertion phase of each patch would rotate the main lobe of the antenna pattern. However, most array designs require element widths of different sizes, which cause unequal insertion phases to occur and modify the pattern in even less desirable ways.

One method to compensate for the insertion phase is to shorten the transmission line lengths between the elements by an amount $\Delta\ell$ such that $\Delta\ell/\lambda_e = \phi_e/2\pi$, where λ_e is the signal wavelength in the dielectric and ϕ_e is the insertion phase. The effect of the dielectric is to change the wavelength dimension from the free space value, λ_0, to the dielectric value, λ_e, by the relation $\lambda_e = \lambda_0/\sqrt{\varepsilon_e}$. The reduction in length by this method causes the patch separation to have the desired equal electrical path lengths, but unequal physical lengths.

Mutual Coupling

The proximity of one patch element positioned close to another in forming an array will cause coupling of the electric (E-) or magnetic (H-) fields, depending on the patch element plane orientation. A series-feed array aligns the E-fields to be oriented along the array direction. Published literature (4,5) and in-house experiments* provide some data for mutual coupling between two nearly square, resonant patches, one driven through a 50-ohm microstrip line and unterminated and the other one terminated in 50 ohms. At a gap width of 0.1 λ_0, the power received by the passive patch is about 17 dB below the power absorbed by the driven patch, and the received power decreases with increasing separation. These data were tabularized in a subroutine and are used in designing an array; however, the effect is normally negligible.

Theoretical Model

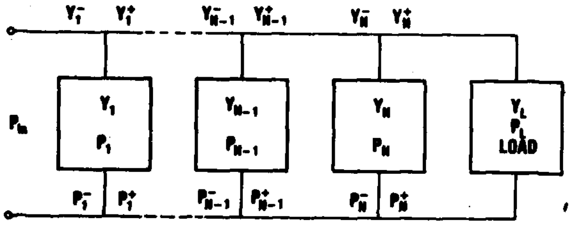
The model for the multielement linear array used in the antenna computer code is represented by a transmission line network with shunt loads as shown in figure 3. Each array element is represented by a normalized shunt admittance, Y_i, that radiates power, P_i. The total admittance of the line to the right of the ith element, seen from this point, is represented by Y_i^+. Y_i^- is the total admittance looking to the right at the ith location, which includes the ith admittance, so $Y_i^- = Y_i + Y_i^+$. The admittance transformation by a line section of length ℓ is given by

$$Y_i^+ = \frac{Y_{i+1}^- \cosh \gamma\ell + \sinh \gamma\ell}{\cosh \gamma\ell + Y_{i+1}^- \sinh \gamma\ell} \quad , \qquad (7)$$

here, $\gamma = \alpha + j\beta$ is the complex propagation constant of the unloaded line, where γ is the complex propagation constant, α is the attenuation constant, and β is the phase constant. The length ℓ includes the electrical length of the patch as well as the interconnecting transmission line.

---

*F. Farrar, Harry Diamond Laboratories.



V_1^- V_1^+
V_{N-1}^- V_{N-1}^+
V_N^- V_N^+
P_{in}
Y_1
P_1
Y_{N-1}
P_{N-1}
Y_N
P_N
Y_L
P_L
LOAD
P_1^- P_1^+
P_{N-1}^- P_{N-1}^+
P_N^- P_N^+

Fig. 3. Equivalent network representation of multielement array model used for computer aided design.

A single-cell analysis of the array elements is illustrated in figure 4a. P_i^- is the power incident to the ith element. If Γ_i^2 is the power reflection coefficient at the ith element such that $P_{(reflected)} = \Gamma_i^2 P_i^-$, then the total net transported power at a point to the left of element i is

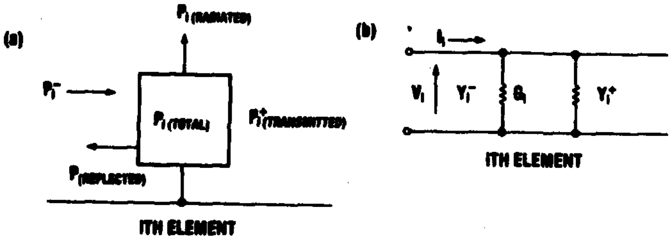
$$\begin{aligned} P_{i(total)} &= P_i^- - P_{i(reflected)} \\ &= P_i^-\left(1 - \Gamma_i^2\right) \quad , \end{aligned} \tag{8}$$

and the power transmitted past the element is the difference between that and the power P_i radiated by the element:

$$\begin{aligned} P_i^+ &= P_{i(total)} - P_i \\ &= P_i^-\left(1 - \Gamma_i^2\right) - P_i \quad . \end{aligned} \tag{9}$$

The patch admittance, Y_i, is actually the aforementioned lumped radiation conductance, G_i (fig. 4b). The radiated power, P_i, may be expressed by G_i and an equivalent element voltage, V_i, as $P_i = |V_i|^2 G_i$. Similarly, the transmitted power is

$$P_i^+ = |V_i|^2 R_e\left(Y_i^+\right) \quad . \tag{10}$$



(a)
$P_{i\,(RADIATED)}$
P_i^-
$P_{i\,(TOTAL)}$
$P_i^+{}_{(TRANSMITTED)}$
$P_{(REFLECTED)}$
ITH ELEMENT
(b)
I_i
V_i
Y_i^-
G_i
Y_i^+
ITH ELEMENT

Fig. 4. ith element: (a) single cell analyis of power distribution and (b) schematic representation.

Note that a reflection from the following patch, $P_{i+1(\text{reflected})}$, is included in P_i^+ because Y_i^+ according to equation (7) is the precisely transformed input admittance of this patch. The total input conductance is

$$R_e\left(Y_i^-\right) = G_i + R_e\left(Y_i^+\right) \quad . \tag{11}$$

With equations (9) to (11), the radiated power can be expressed as

$$P_{i(\text{radiated})} = P_i^-\left(1 - \Gamma_i^2\right)G_i/R_e\left(Y_i^-\right) \quad , \tag{12}$$

where

$$\Gamma_i = \left|\frac{1 - Y_i^-}{1 + Y_i^-}\right| \quad .$$

Equation (12) can be used to determine the power radiated from any element in the array, provided that one other relationship is known. At the first element, P_1 is the input power (usually normalized to 1). The total input admittance is Y_1^-, and G_1 is the first element conductance. To transform the power from the first to the second element, an additional relationship,

$$P_{i+1}^- = P_i^+ e^{-2\alpha\ell} \quad , \tag{13}$$

needs to be used* so that from equation (9)

$$P_2^- = P_1^+ e^{-2\alpha\ell} = \left[P_1^-\left(1 - \Gamma_1^2\right) - P_1{*}e^{-2\alpha\ell} \tag{14}$$

and

$$P_2 = P_2^-\left(1 - \Gamma_2^2\right)G_2/R_e\left(Y_2^-\right) \quad . \tag{15}$$

The above equations appear as recursion relations, which are quite useful for generating and analyzing N-element array designs with the aid of a computer. Given the normalized conductor elements, G_i, the spacial element separation, ℓ, and the propagation coefficients, α and β, the normalized power radiated from each element is determined. However, array designs for antennas usually prescribe the radiated power from each element as input data, and the engineer must determine the element conductances that form the array--in essence, working the problem backward. This task requires the solution to an implicit equation, that is, having knowledge of the conductance values to determine the conductance values. To solve the task, a computer program was written by using the explicit relations of $P_i = f(g_i)$ with an optimization subroutine.* This subroutine continuously alters the conductance value of each element until it finds a set of values such that the calculated power best approximates in a least squares sense the prescribed radiated power distribution.

---

*Equation (13) relates the exponentional decay of the power from losses as it translates along the transmission line from one element to the next.

Computer Model

The computer program reported here was developed to provide design data for fabricating a linear series microstrip array antenna. The design is based on information provided by the user as he answers a series of questions that are displayed on the terminal screen.

The program computations normalize impedances or admittances to the line characteristic impedance or admittance as chosen. Distance also is dimensioned by normalizing to the wavelength of interest, as it is measured in the dielectric or free space if so chosen. The element separation need be dimensioned only as decimal parts of a wavelength. Actual design dimensions need not be determined until the antenna design is completed and resonant frequency is assigned.

The program also offers options for considering the effects of (a) reflections due to impedance mismatch at each element, (b) patch directivities due to the broadside gain of each element, (c) insertion phase, and (d) compensation for insertion phase. This capability and the others mentioned provide for a variety of parametric studies toward determining the casual relations of the antenna pattern peculiarities and sensitivity studies important in understanding the effect of manufacturing tolerances.

If the conductance values of the array elements are unknown, they are calculated by the program after the power distribution in entered. Six commonly used power distributions are programmed and are available for use. Other power distributions can be entered manually at the keyboard.

Calculation of Conductances

Equation (12) can be rewritten so that

$$G_i = \frac{P_{i(\text{radiated})}}{P_i^-\left(1 - \Gamma_i^2\right)} R_e\left(Y_i^-\right) \quad . \tag{16}$$

---

*Copyright 1978 by IMSL, Inc.

This equation is not explicit in G_i since Γ_i^2 and Y_i^- also contain G_i. However, an optimum design requires minimum or no reflections $(\Gamma_i^2 << 1)$ along the array elements resulting from good impedance matching $(R_e Y_i^- \simeq 1)$. Therefore, as a first estimation of G_i, Γ_i^2 and Y_i^- are removed and the relation $G_i = P_{i(\text{radiated})}/P_i^-$ is used. Once the set of G's is calculated from the given power distribution, the G's are then used to recalculate the radiated power distribution and the antenna efficiency by use of equation (12). The efficiency of an array used in this report is defined as the ratio of the total power radiated to the power delivered at the input. (Losses include the power absorbed at the termination and the power reflected from each element back to the generator.) For low efficiency design, the power distribution calculated is in fair agreement with the initial power distribution used as input data to the program. Figure 5 plots the initial and given power distributions that were calculated to radiate from the elements. Better fits can be obtained by using the optimization subroutine.

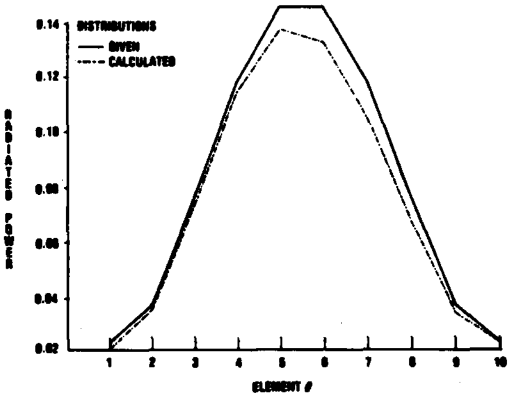


Fig. 5. Initial and given power distributions calculated to radiate from elements.

Optimization

An algorithm in the program computes an initial estimate of the required element conductances. However, for most applications of the array code, the resulting low efficiency or undesired distribution of radiated power makes the code somewhat limited or useless. An optimization routine can be used as an option to provide a better or perfect match of the two power distributions. This routine calls on a subroutine that generates the calculated power distribution from the calculated radiated power of each element and compares the distribution with that given as input data. The square of the difference is calculated, and the optimization routine varies the conductance values to minimize this difference. When the calculations are completed, the power calculated to radiate from each element is in very good agreement with the input power data, and the conductance values generated are used for designing the array. Figure 6 shows an example of the results of the optimization routine used for a particular design.

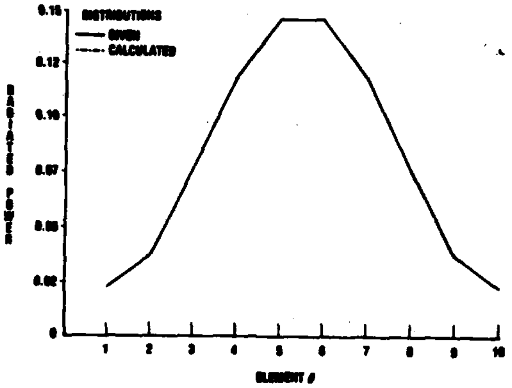


Fig. 6. Results of optimization routine used for particular design.

The antenna radiation pattern from the generated design can be displayed on the terminal screen over any desired angular range for either linear or polar plots.

Power Distribution Subroutines

The power distribution to be radiated over the antenna aperture may be inserted into the program either by use of a terminal keyboard or by use of the following distributions already available in the program as subroutines:

- Uniform
- Tchebyscheff
- Taylor
- Cosine square on pedestal
- Binomial
- $(1 - x^2)^m$ polynomial

The choice of any specific distribution must include tradeoffs among gain, beam width, and side-lobe levels, a choice usually dictated by system requirements. Here, the computer code is particularly useful in displaying the design for each distribution, where comparisons of each feature may be explored with relative ease. However, it is important to have some basic knowledge of the various distributions before attempting a design. The uniform distribution provides for the highest gain at a narrow beam width, but the side lobes are high (-13.2 dB).

The Tchebyscheff and Taylor distributions provide for side-lobe control at some loss to antenna gain. The Tchebyscheff distribution produces the narrowest beam width for a given side-lobe level; however, the side lobes are equal. For many applications, the equal-side-lobe-level pattern may be undesirable. In the Taylor distribution, the side lobes decrease as the angle from the main beam increases, but at the sacrifice of a slightly increased beam width. Both the Tchebyscheff and

Taylor distributions enable the designer to select a desired side-lobe level and to compute the element excitation values in terms of a parameter related to the side-lobe level.

The distribution of cosine square on a pedestal is a truncated distribution in which the ratio of the pedestal height to the distribution peak determines the side-lobe level. The binomial distribution produces no side lobes, but at the expense of low gain and wide beam widths. The polynomial distribution provides for a class of distributions in which further tradeoffs among gain, beam widths, and side lobes can be studied simply by varying the factor m in the polynomial.

The variety in aperture distributions offered by this program should provide for most systems needs. New or unique distributions can be added easily to the existing program as subroutines to expand the scope of this design code.

Antenna Designs

Several linear and planar arrays were designed by using the microstrip computer model. Figure 7 shows a microstrip S-band array that was successfully designed to exhibit the same antenna characteristics (except for polarization) as those derived from the slotted waveguide array. The microstrip array, however, is E-plane polarized along the array direction, and the waveguide array is H-plane polarized. A number of linear arrays were combined by use of a corporate-feed structure to form a planar array as shown in figure 8. Figure 9 shows a polar plot of the antenna radiation pattern for both the E-plane (series-feed direction) and the H-plane (corporate-feed direction) polarizations. The success at designing microstrip antennas at microwave frequencies prompted an array design for use at around 100 GHz. Figure 10 shows the array mounted in a test fixture, and figure 11 shows the linear plot for both the theory and experimental results.

Fig. 7. Microstrip antenna designed to have characteristics similar to those of waveguide array shown above.

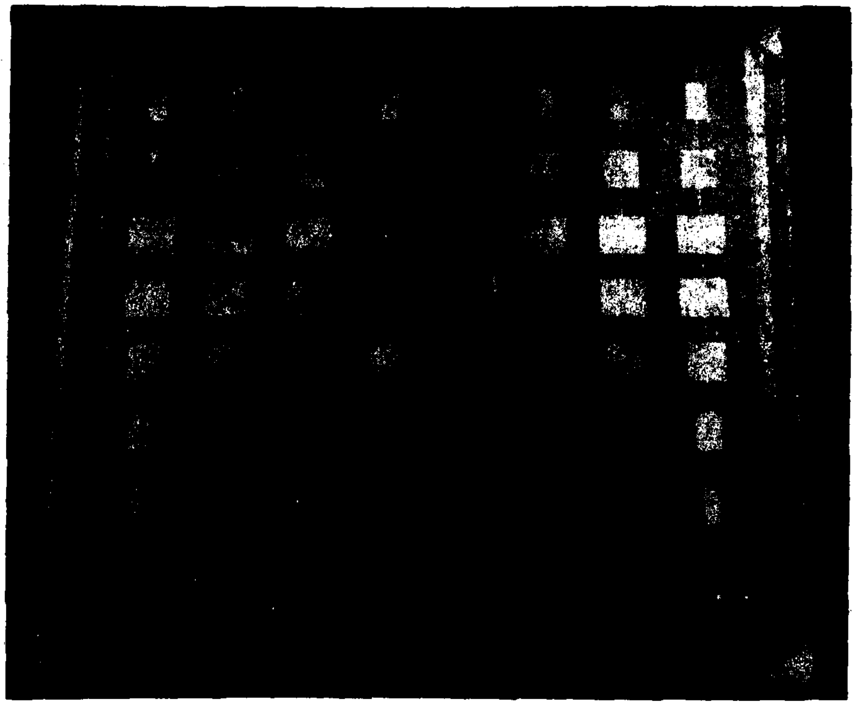


Fig. 8. Microstrip planar array.

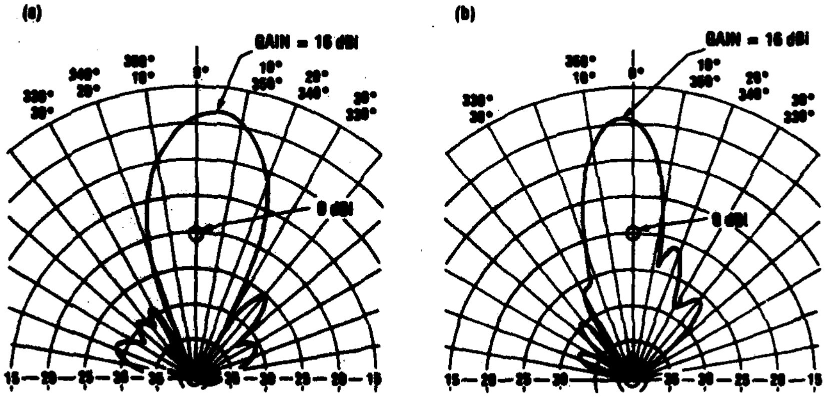


Fig. 9. Polar plots of antenna radiation patterns: (a) E-plane and (b) H-plane.

Fig. 10. Microstrip design of first U.S. Army near-millimeter wave antenna.

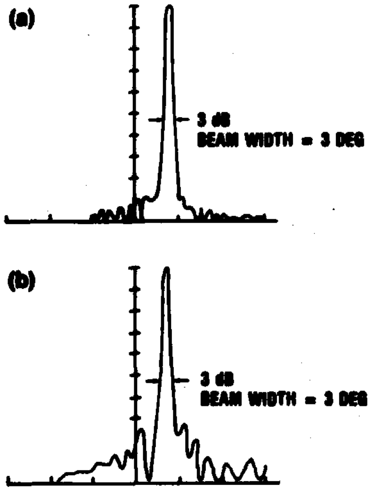


Fig. 11. Radiation patterns of first U.S. Army near-millimeter wave monolithic conformal antenna with linear phased array: (a) theory and (b) experiment.

Conclusions and Recommendations

By using a simple transmission line model of shunt conductance elements along with present computer programming techniques, linear array microstrip antennas can now be simulated, designed, and studied in a short time and at a low cost. Measurements of resonant frequency and beam angle of all the antennas fabricated show excellent correlation with the design data submitted to the computer. Cross polarization of the fields and mutual coupling of the array elements appear to be minimal. Several antennas etched by using the same negative have shown identical return-loss measurements and radiation patterns indicating excellent reproducibility.

Additional work is needed in obtaining a better analytical model and data on the element values of conductance and insertion phase as a function of patch width, characteristic impedance, and loading. When incorporated with an ABCD parameter and transmission matrix calculations, this approach should improve the antenna side-lobe levels and the overall measured design performance. Additional information must be gathered to better describe the radiation pattern and the directivity of a single patch element. Further, it would be desirable to design series arrays having elements of equal widths, but with different characteristic impedance lines interspaced between them. This design would provide for a simpler means of phase shifting the array since the insertion phase of each element is the same. Additionally, information about thick and low-loss substrates is needed so that arrays may be better designed to 100 GHz, where most of the commonly used dielectrics have thicknesses approaching the dimensions of one wavelength.

References

(1) A. G. Derneryd IEEE Trans. Antennas Propag. AP-24, Nov 1976, 846-851.

(2) I. G. Bahl and P. Bhartia, Microstip Antennas, Artech House, Inc., 1980.

(3) T. Metzler, Proc. Workshop on Printed Circuit Antenna Technology, New Mexico State Univ., Las Cruces, Oct 1979, pp 20/1-16.

(4) P. P. Jedlicka and K. R. Carver, Proc. Workshop on Printed Circuit Antenna Technology, New Mexico State Univ., Las Cruces, Oct 1979, pp 4/1-19.

(5) P. P. Jedlicka and K. R. Carver, IEEE Trans. Antennas Propag. AP-29, Jan 1981, 147-149.

Resolution of Factors Responsible for Difficulty in Growing Single Crystals of YAG

*Dr. Jaroslav L. Caslavsky and Dr. Dennis J. Viechnicki
Army Materials and Mechanics Research Center
Watertown, Massachusetts 02172

INTRODUCTION:

In a ternary system like Al_2O_3-Y_2O_3-Nd_2O_3 it is naturally realistic to expect that numerous factors will contribute, each in its own way, to the complexity of growth problems of Nd:YAG single crystals which is a compound in the ternary system. However, extensive experimental study indicates that not many factors but rather a single factor has caused a derogative influence on the growth of Nd:YAG single crystals.

DEFINITION OF THE PROBLEM:

In 1980, Caslavsky and Viechnicki (1) reported that after melting of YAG, the structure of the melt undergoes a change, i.e., in spite of congruent melting of YAG there is no equilibrium between the solid and molten YAG. The non-equilibrium condition occurs on account of coordination change of 3/5 of the aluminum ions, which in the solid YAG structure are present in four-fold coordination, while in the melt they increase their coordination to six. On cooling such a melt, in which all YAG nuclei were destroyed, the coordination change is irreversible and the melt of YAG composition begins to solidify at 1855°C where yttrium aluminate (perovskite) commences to solidify as a metastable proeutectic phase. The solidification ends at 1700°C as a metastable eutectic mixture of 23.0 mol% Y_2O_3 and 77.0 mol% Al_2O_3. On the other hand, if a YAG nucleus, say a seed crystal, is present in the melt and latent heat of solidification is slowly withdrawn from the solid-liquid interface through the seed crystal, the YAG structure will perpetuate from the seed crystal on account of the melt converting into a YAG single crystal. It also was reported if the rate of solidification increases over a certain rate; precipitation of metastable proeutectic phase occurs and the amount formed is directly proportional to the rate of solidification. Further, it was suggested that presence of neodymia in the melt, as a dopant, somehow de-stabilizes four-fold coordination of the aluminum ions, therefore, the amount of second phase in the presence of neodymia increases with the rate of crystallization more rapidly than in the undoped melt.

To understand this phenomena, futher experimental evidence to explain the enhanced formation of scattering centers in doped YAG crystals was sought and is described in this paper.

EXPERIMENTAL METHODS:

a) Growth of the Single Crystals by the Vertical Solidification of the Melt (VSOM):

In this technique a cylindrical crucible of approximately 75 mm diameter and 125 mm height is used as a container. The seed crystal is placed in the middle of the bottom of the crucible which then is filled with the material to be melted. The crucible is positioned with its center on a water cooled heat sink, which in addition to serving as a support for the crucible, prevents the seed crystal from being melted and provides the means for controlled latent heat removal of solidification during the period of the crystal growth. Since the heat sink is now in the exact center of the cylindrical heating element, the crucible is heated evenly in the circumference, but there is an upward thermal gradient built into the element which is consequently reflected in the melt. Unlike most other crystal growth techniques there are no moving parts in the VSOM crystal growth technique. The driving force for crystallization in this technique is derived from the motion of the thermal field throughout the hot zone of the furnace. Elimination of moving parts combined with the upward thermal gradient in the melt gives to the VSOM technique a unique stabilization condition, i.e., there is practically no thermally induced flow inside the melt, and mechanically produced vibrations are absent.

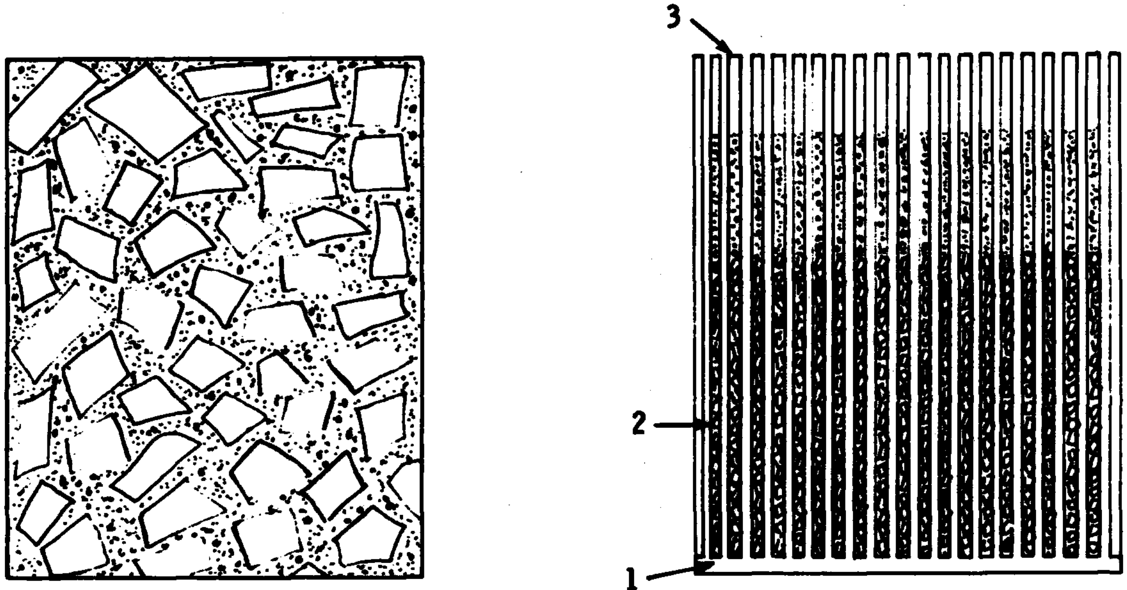
b) The Optical Differential Thermal Analyses (ODTA):

This technique was well described in (1) and since during the course of this investigation it was used only as a supporting method the reader is referred to the pertinent publication.

c) Experimental technique to reveal the influence of $NdAlO_3$ on the Solidification of Nd:YAG Melt:

From the figure 1a and 1b it is apparent that when the growth parameters are well controlled, the difference in the amount of scattering centers in single crystals of YAG and Nd:YAG grown by VSOM method are in both cases minimized or absent. Therefore, to study the factors responsible for the formation of second phase particles from single crystals is difficult. Owing to this fact it was decided to study the differences of solidification of unseeded doped and undoped melts of YAG composition.

Except for the absence of the seed crystal, the solidification process was exactly the same like it is used by growing single crystals by VSOM technique. The 75 x 125 mm molybdenium crucible was filled with material prepared by a process developed for the preparation of charges for growing YAG crystals as described in (2). The crucible was placed in VSOM crystal growth furnace and heated at a rate of 50°C/h



a. YAG

b. Nd : YAG

1. Single Crystal of $YAlO_3$
2. $YAlO_3$ + Al_2O_3 Crystals
3. $YAlO_3$ Dendrite

Figure 1. MELTS SOLIDIFIED WITHOUT THE SEED CRYSTAL.

until 1940°, the melting point of YAG, was reached as indicated by a thermo-arrest on the ODTA. At this point, the temperature increase was lowered to 5°C/h and after the thermo-arrest ended, about 90 minutes later, the temperature of the melt was increased to 2050°C. The melt was kept at this temperature for 18 hours to insure that no solid particles were present within the melt. Then the melt was cooled down at a rate 0.3°C/h until the entire content of the crucible was solidified as measured by the ODTA.

RESULTS:

a) Solidification of the undoped melt. During the temperature decrease the ODTA zero line remained stable until 1970°C, at which point a very sharp large exothermic maximum revealed spontaneous nucleation in the entire mass of the melt contained in the crucible. The solidification material removed from the crucible was in the form of large $YAlO_3$ crystals interspersed with much finer Al_2O_3 crystals. The overall appearance of the material was similar to a rapidly solidified mixture of high melting oxides (see fig. 1a).

b) Solidification of the doped melt. Material of composition $Y_{2.97}Nd_{0.03}Al_5O_{12}$ (corresponding to 1at% Nd:YAG single crystals) was melted and solidified as described for the undoped melt. However, significant differences were observed by ODTA solidification of the doped melt as compared to the undoped melt. At 1825°C an exothermic maximum was observed similar in character and magnitude to one which indicates freezing of the last liquid, during the regular crystal growth. The last liquid during crystal growth freezes at 1925°C since it is richer in yttria, due to the incongruent evaporation of alumina. The solidification of the unseeded melt follows the metastable crystallization path, hence, the exothermic maximum at 1825°C indicates freezing of a liquid richer in the Al_2O_3 than the YAG melt. With the further temperature decrease at 1700°C a second small exothermic maximum is observed due to the freezing of the metastable eutectic; this maximum is never obseved during the crystal growth process.

The appearance of a solid ingot obtained by solidification of an unseeded doped YAG composition melt is quite different to the undoped one. The entire bottom of 75 mm diameter solid ingot was transparent to the height of about 5 mm. The x-ray analysis revealed that it was a single crystal of $YAlO_3$ with the (110) crystal plane parallel with the bottom of the crucible. Since the $YAlO_3$ single crystal grew from the melt of different composition, the constitutional supercooling consequently instituted dendritic growth. The perovskite dendrites continued to grow until the composition of the metastable eutectic (23.0 mol% Y_2O_3 and 77.0 mol% Al_2O_3) was reached at 1700°C. At this point the metastable eutectic mixture solidified. In consequence of 22% of negative volume change, due to the liquid-solid transformation of the eutectic liquid, the level rapidly descended which revealed a "forest" of perovskite dendrites all parallel to each other with the dendritic axis in the [001] direction. This indicated that all of the densities were seededby the single crystal found on the bottom of the ingot. The metastable eutectic mixture was found in the spaces among the dendrites (see fig. 1b).

These experiments were repeated three times with the same results; the reproducibility of the experiments and their interpretation led to a mechanism of scattering centers formation in Nd:YAG single crystals.

EXPLANATION OF FORMATION OF SCATTERING CENTERS:

In the solid state the YAG structure, from the oxides components Al_2O_3 and Y_2O_3, forms quite easily at temperatures as low as 700°C. The kinetics of formation of the YAG structure is directly proportional to the temperature, and at the temperatures between 1700°C and 1800°C the stoichiometric mixture of the oxides reacts completely in two hours to the YAG structure. By contrast, when the oxide mixture is melted and then solidified no YAG is formed. This phenomenon is seen whether or not Nd_2O_3 is present. Nevertheless, if the melt is seeded and solidification performed under the controlled conditions, a single crystal of YAG is obtained. However, the experimentally determined growth rates needed to obtain scattering center free doped crystals, are at least three times lower than the growth rates for undoped crystals.

In reference (1) it was established that it was impossible to form YAG from an unseeded melt due to the fact that the aluminium ions are not willing to lower their coordination without external influence and will much rather form incongruently melting perovskite even from non--stoichiometric melt. The only mechanism which forces aluminum ions to the four-fold coordination is the dangling bonds protruding from the free surface of the seed crystal. Such dangling bonds in the most favored case can extend from the solid structure of the YAG only 12A. Therefore, the coordination change can take place only in close proximity to the free surface of the seed crystal. If the solidification rate exceeds the rate with which the coordination change can take place, the perovskite phase will crystallize simultaneouslywith the YAG. From all experimental work it is evident that the rate of change of the aluminum ions is the factor controlling the rate of the YAG structure formation. If this would be the only factor responsible for formation of the scattering centers in YAG, there should be no reason for Nd:YAG to grow scattering center free only at much lower rates than undoped YAG. From the experiments described in the previous paragraph, the following conclusion for enhanced formation of scattering centers in Nd:YAG is interpreted as follows:

$YAlO_3$-perovskite is the highest melting compound in the ternary system melting congruently without decomposition at 2110°C. It is highly probable in this ternary system that under sufficient supercooling $NdAlO_3$ will first form stable or supercritical nuclei, and this fact accounts for the difference in the crystallization behavior of unseeded doped and undoped melts of YAG composition.

In the VSOM crystal growth arrangement the melt around the bottom part of the crucible nearest the heat sink is estimated to be cooler by 100° to 150°C than the melt farther away from the heat sink. During slow cooling of a melt of Nd:YAG composition there is enough time, in the locally supercooled area, for $NdAlO_3$ to nucleate and eventually form supercritical nuclei. With progressing cooling such nuclei will grow. However, in this melt there is not enough neodymium ions to sustain the growth of $NdAlO_3$ but since $NdAlO_3$ is istostructural with $YalO_3$ a heteroepitaxial boundary will be formed and $YAlO_3$ will continue to grow off as a single crystal until the constitutional supercooling will force dendritic growth.

In the VSOM technique it is impossible experimentally to measure the thermal gradient near the solid-liquid interface, but from experience gained working with this technique, it has been found that to control growth, the thermal gradient near the interface has to be very steep. In such steep gradient, however, nucleation of the $NdAlO_3$ still can occur according to a reaction

$$L_1 = L_2 + S \quad \ldots\ldots \quad (1)$$

where L_1 is a liquid free of any solid particles while L_2 is liquid whose composition has changed due to the nucleation of $S = NdAlO_3$. Therefore, if the rate of the growth of Nd:YAG exceeds the transportrate of solid particles so that they cannot be carried away from the solid-liquid interface into the melt region where temperature is high enough to reverse reaction (1) to:

$$L_2 + S = L_1 \quad \ldots\ldots \quad (2)$$

then the $NdAlO_3$ particles will deposit by gravity on the solid part of the solid-liquid interface and will be embedded in it and will form scattering centers in the resulting single crystal.

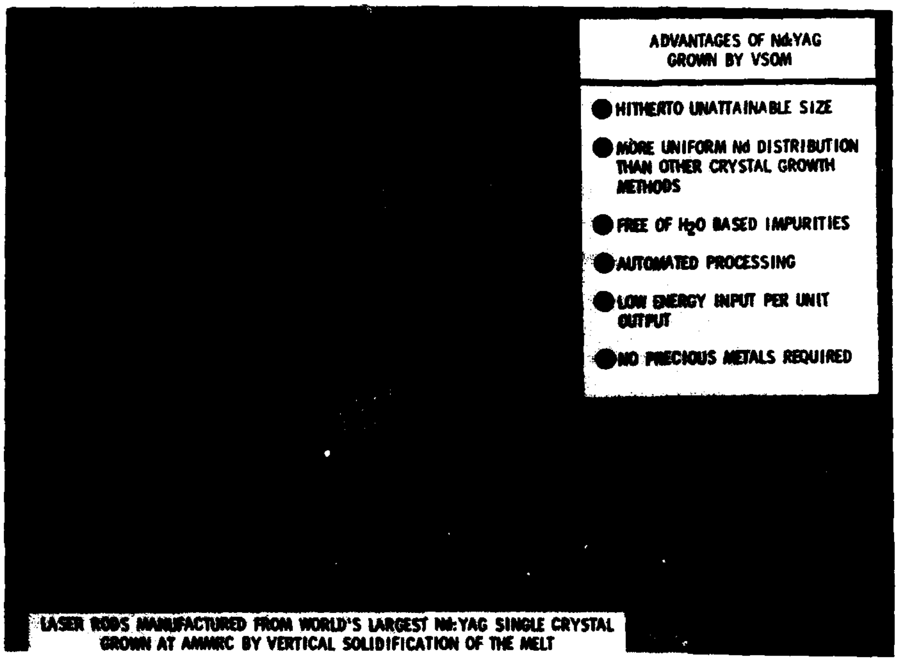
CONCLUSIONS:

The enhanced formation of the scattering centers in Nd:YAG is attributed to the fact that the melting point of YAG is 170°C lower than the melting point of $NdAlO_3$. Further, since aluminium ions in the neodyium aluminate are present in six-fold coordination $NdAlO_3$ can easily nucleate near the solid-liquide interface where lower temperature favors reaction (1). Provided that the growth rate of Nd:YAG exceeds the kinetics of reaction (2), the second phase particles will precipitate in the growing Nd:YAG crystal and appear as scattering centers.

From the experimental evidence it can be concluded that in the composition section on and near the tie line connecting $NdAlO_3$ and $Y_3Al_5O_{12}$ in the Nd_2O_3-Y_2O_3-Al_2O_3 ternary system, $NdAlO_3$ shows an inherent property to nucleate first. To eliminate the nucleation of $NdAlO_3$ during the crystal growth of Nd:YAG the thermal gradient near the solid-liquid interface has to be very steep while the melt temperature has to be kept above the temperature which prevents reaction (1). Simultaneously the speed of the interface advancement has to be set below the thermally induced flow of the melt from the solid-liquid interface. Therefore, it is apparent that with the correct growth conditions scattering center free Nd:YAG single crystals can be grown by VSOM.

REFERENCES:

1. Caslavsky, J.L., Viechnicki, D.J., Journal of Material Science 15 (1980) 1709.
2. Viechnicki, D.J., Caslavsky, J.L., American Ceramic Society Bulletin 58 (1979) 700.



ADVANTAGES OF Nd:YAG GROWN BY VSOM

- ● HITHERTO UNATTAINABLE SIZE
- ● MORE UNIFORM Nd DISTRIBUTION THAN OTHER CRYSTAL GROWTH METHODS
- ● FREE OF H_2O BASED IMPURITIES
- ● AUTOMATED PROCESSING
- ● LOW ENERGY INPUT PER UNIT OUTPUT
- ● NO PRECIOUS METALS REQUIRED

LASER RODS MANUFACTURED FROM WORLD'S LARGEST Nd:YAG SINGLE CRYSTAL GROWN AT AMMRC BY VERTICAL SOLIDIFICATION OF THE MELT

Figure 2

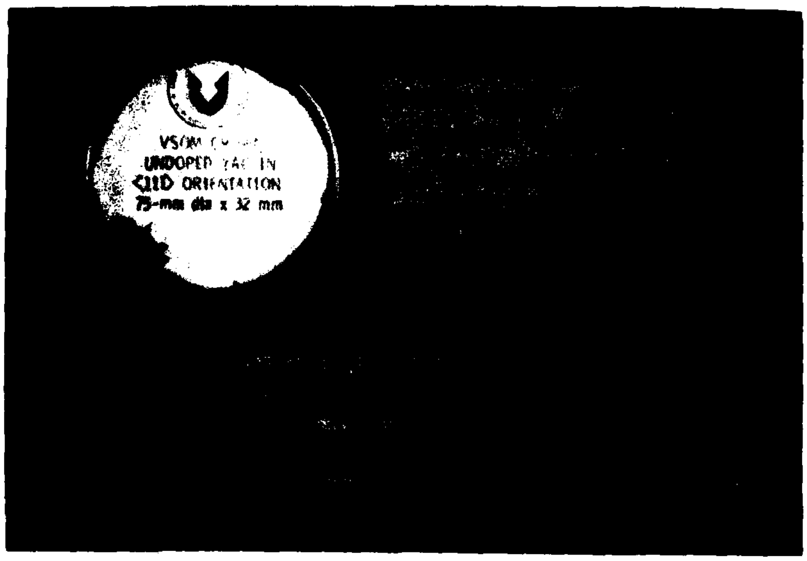

VSOM GROWN
UNDOPED YAG IN
<111> ORIENTATION
75-mm dia x 32 mm

Figure 3

GENERATION OF A SYSTEM WHICH EXIBITS AN ISOPYCNIC IN THE CONVERSION OF ARMY VEHICLES TO SILICONE BRAKE FLUID (1)

*MR. CHARLES C. CHAPIN
MR. JAMES H. CONLEY
MR. ROBERT G. JAMISON
US ARMY MOBILITY EQUIPMENT RESEARCH AND DEVELOPMENT COMMAND
FORT BELVOIR, VA 22060

The US Army has been using three different automotive hydraulic brake fluids covered by Federal Specification VV-B-680, for use in all Tank-Automotive equipment (temperate -tropical areas), Military Specification MIL-H 13910 for arctic use, and Military Specification MIL-P-46046 for preservative use (2). These polyglycol and castor oil type fluids are hygroscopic and absorb water while in use which adversely affects their performance by lowering the vapor lock temperature, increasing the low temperature viscosity, and by contributing to component corrosion which can lead to cup scoring, as well as system failure.

In 1967 the US Army began developing a single multipurpose silicone based brake fluid which would overcome the absorption of water exibited by the conventional fluids as well as provide all weather and preservative properties. This one fluids, which replaces the three existing fluids, can reduce logistics and maintanence costs. Any vehicle equipped with this fluid will be ready for use in any geographical area at a moments notice. Therefore, vehicles could be moved from the tropics to the arctic or could be put into storage without requiring any other current brake system maintanence procedures (3).

Brake Fluid, Silicone (BFS), Military Specification MIL-B-46176, which was developed by MERADCOM (4) in conjunction with industry was approved for Army use in 1980. All tactical vehicles, administrative use vehicles, commercially procured vehicles, construction equipment and material handling equipment which currently use polyglycol type fluids will be converted to silicone brake fluid by July 1982.

The wipe and clean procedure for conversion, which was successfully tested at Yuma Proving Grounds, Arctic Field Test Center, Panama Tropic Test Center (5) and was the recommended conversion method, was found to be unacceptable due to the labor intensiveness of the procedure, its implications to manpower requirements and cost. An alternate method (6) for brake fluid replacement, the flush and fill procedure (adapted for TB43-

0002-87) was approved for use based on these same considerations. A straight flush and fill procedure was tested by MERADCOM (7) and found to be inefficient for the following interdependent reasons:

a. The geometry of the wheel cylinders (bleeder valves at the top).
b. The immiscibility of the two types of fluid.
c. The lower density of the silicone relative to the polyglycol.

Figure 1 shows a cross-section of each of the wheel cylinder types and a caliper. The inlet lines and bleeder valves are indicated. The bleeder valves are always positioned at the uppermost point of the wheel cylinder to allow the lower density air to be bled from the system.

The results of testing of straight flush/fill procedure indicated that for the M-151 (jeep) vehicles, the procedure removed essentially all of the polyglycol. However, the other wheel cylinders (and calipers) do not have inlet lines located at the bottom of the cylinder and the method leaves residual polyglycol trapped in the system. Since inlet lines are in the middle of the cylinder, the fluid simply overlayers the polyglycol after removing the accessible polyglycol by displacement.

An ILIR project was initiated at MERADCOM to address this problem experimentally and to subsequently develop a cost-effective method which would achieve complete polyglycol removal.

Approach

Since the straight flush/fill method gave a liquid binary phase system (silicone polyglycol) a solution approach was used to reverse the phase structure by the inclusion of an intermediate fluid of sufficiently low density. To achieve a reversal of the phases, a fluid must be selected for which the ternary system will exibit an isopycnic tie line (8). The isopycnic tie line (or twin density tie line as it is called in multicomponent systems) is that tie line at which the densities of the two phases are equal (a zero density difference). Upon passing the isopycnic line the density difference for the two phase undergoes a sign change (plus to minus or visa versa) and the phases are reversed, the upper becoming the lower and the lower becoming the upper.

The approach was, then to select a fluid which would generate an isopycnic in the silicone-polyglycol system with as high a phase boundary as possible (so that two phases will form under the various conditions found in the different wheel cylinders and calipers). This intermediate fluid must also be a cosolvent for the silicones and the polyglycols, have a high boiling point, be non-corrosive, non-hygroscopic, non-toxic, non-flammable, be inexpensive, and must be compatible with the elastomers in brake systems.

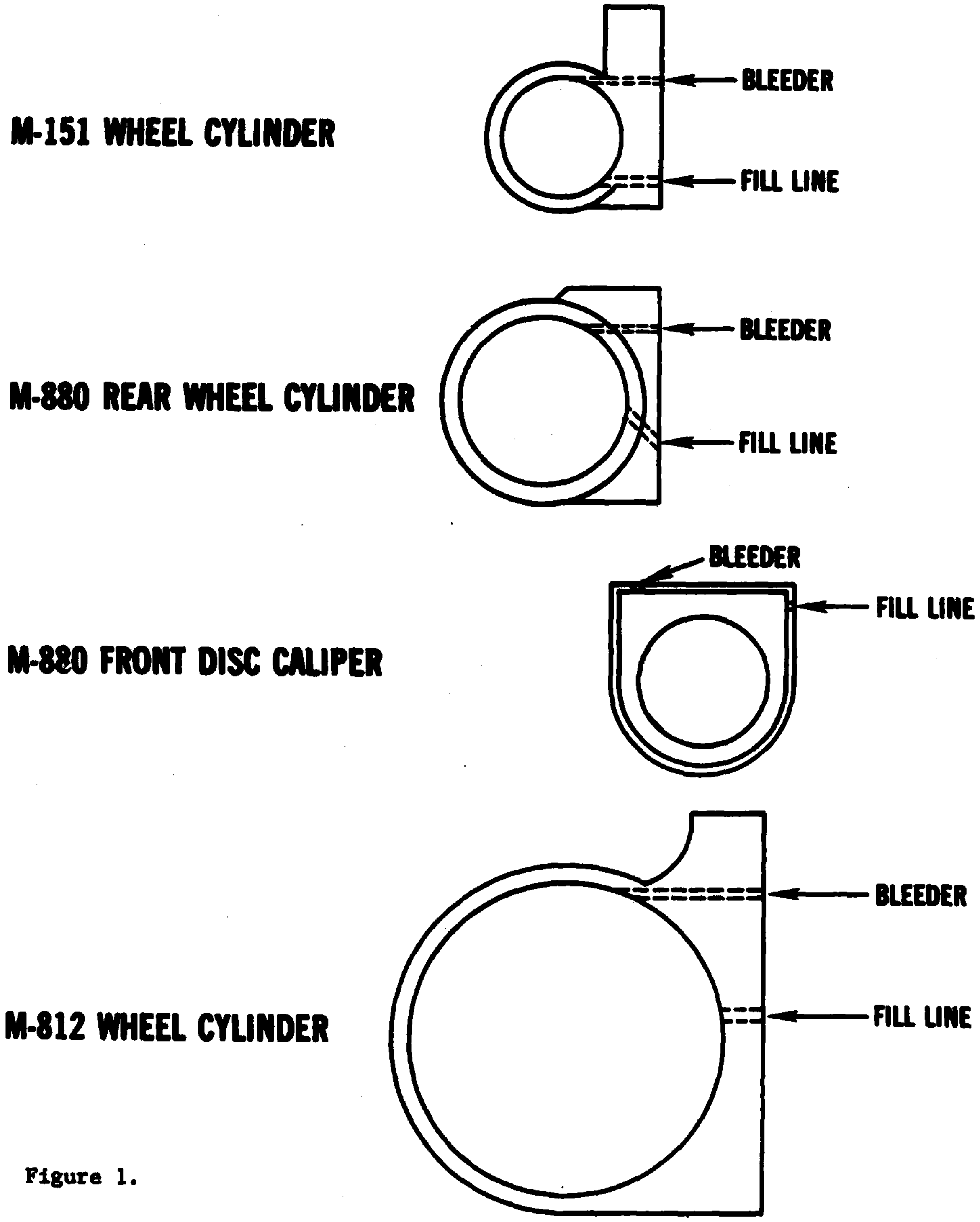

M-151 WHEEL CYLINDER
BLEEDER
FILL LINE
M-880 REAR WHEEL CYLINDER
BLEEDER
FILL LINE
M-880 FRONT DISC CALIPER
BLEEDER
FILL LINE
M-812 WHEEL CYLINDER
BLEEDER
FILL LINE

Figure 1.

Description of Tests

Air Flushing.

a. Apparatus. A wheel cylinder from a 5 ton truck with a spring, cups, and pistons was mounted in a vise and copper tubing was connected from the inlet to a syringe which had a three way valve attached to it. The syringe allowed filling of the wheel cylinder from a reservoir. A vent tube was attached to the bleeder valve of the wheel cylinder and directed to the drain.

b. Experimental. This system was used to assess the utility of air flushing for polyglycol removal. Water was used initially to check for leaks in the system.

Solvent Addition.

a. Experimental. A solvent (Table 1) known to be miscible (9) with silicones was selected (2-ethyl hexanol, 2-EH) and 1 ml was added to 1 ml of a polyglycol brake fluid. The two were miscible. To this mixture was added 1 ml of BFS. The BFS went to the bottom of the tube and retained its color. Upon vigorous shaking, the mixture formed an emulsion which settled in about 20 minutes into two distinct layers. The top layer was not transparent (a possible microemulsion) but after sitting overnight, it was clear. Gas chromatographic analysis of the layers revealed that the bulk of the polyglycol and the 2-EH were in the yellowish upper phase, and that a small portion was in the lower clear phase (the dye from the BFS went into the polyglycol layer). The 2-EH was then tested and found to be a solvent for BFS (Figure 2).

Wheel Cylinders And Plexiglass Windows.

a. Apparatus. For the purpose of visually observing the mixing, a wheel cylinder from a 5 ton truck was equipped with end plates made of plexiglass and bolted into place to provide a fluid seal. The spring, pistons and cups were removed for viewing, and the dust boots were trimmed so that the window diameter was approximately 1-½ inches for easy viewing of the process. A lamp was placed at the rear of the assembly so that the interior of the cylinder was illuminated.

b. Experiments. The wheel cylinder was charged with polyglycol and a series of experiments was performed for the purpose of determining the feasibility of the method and ascertaining if effective mixing could be accomplished. The degree of mixing, aeration of the fluids and phase behavior were observed visually.

Mock-Up of A Brake System.

a. Apparatus. A system of six wheel cylinders and a master cylinder was constructed, and they were connected with ¼ inch cooper tubing. The wheel cylinders were bolted to a bracket so that they could be bled properly.

b. Experiments. Systems of long and short lines were used, and

Table 1. Properties Of 2-Ethyl Hexanol (2-EH).

Molecular Weight	130.22
Boiling Point	364.6°F
Freezing Point	-104.8°F
Density(g/cm^3)	0.8323
Vapor Pressure (68°F, mbar)	0.5
Water Absorption (%)	2.7
Flash Point	178.°F
Ignition Temperature	482.°F

Table 2. Residual Polyglycol After Air Flushing.

	ml	%
Air Flush	51	37
Reverse Flush	8	9.4
Vacuum	49	57.6

Note: The volume of the wheel cylinder was 85 ml.

Table 3. Results Using Wheel Cylinders And Plexiglass Windows.

	Extent Of Mixing	Aeration Of Contents	BFS On/Top Bottom	BFS Aerated	Mixing Of BFS/Glycol
Add 2-EH By Syringe	None	-	-	-	-
Air Purge Of Glycol then Hexanol	None	Aerated Polyglycol	-	-	-
Air Purge Before And After Hexanol	Complete	Aerated Polyglycol	BFS/Top Aerated Glycol/ Bottom	No	None
Pressure Flushing 2-EH	Incomplete At Bottom	-	-	-	-

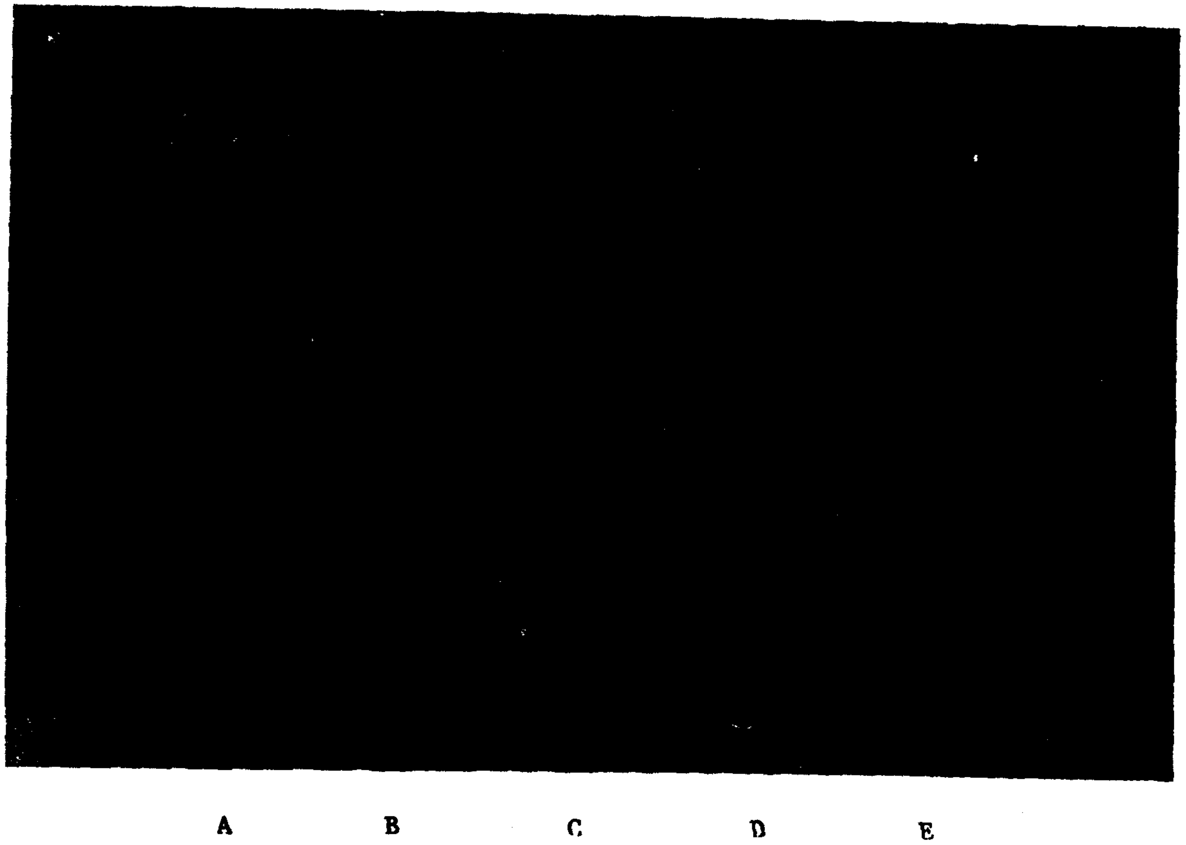


Figure 2. The Phase Inversion Phenomenon. A= polyglycol fluid; B= silicone fluid; C= after addition of silicone to polyglycol, silicone on top; D= 2-ethyl hexanol (2-EH); E= after mixing polyglycol and 2-EH and then adding silicone, silicone on the bottom.

different flushing volumes of the density modifier were used to develop a method for eventual use in the conversion of vehicles.

Caliper Experiments.

a. Apparatus. A caliper from a commercial vehicle was attached to the master cylinder with a three foot line, and spacers were clamped into the caliper to fix the internal volume at 90 ml.

Used Wheel Cylinder.

a. Apparatus. A used wheel cylinder from a 5 ton truck was connected to the master cylinder, and flushing experiments were conducted to determine the effect of the sludge and the extent of sludge removal for this process.

Administrative Vehicle Test.

a. Apparatus. This vehicle was selected because it had wheel cylinders and calipers, and because it had over 100,000 miles on it so it could be expected to be a worst case test.

b. Test. A proposed procedure based on previous experiments was used and the vehicle converted. This procedure involved purging the system with air, flushing the two rear wheel cylinders simultaneously and then flushing with the 2-EH. The wheel cylinders were flushed with one volume of the master cylinder (200 ml) and the calipers were flushed individually with twice this amount of fluid. After the fluid flushes, the air (used for flushing) was allowed to purge the system for one minute after the fluid was purged. The vehicle was then pressure bled with BFS while allowing a specific volume of the BFS to pass through the system to insure complete removal of the 2-EH. The wheel cylinders and calipers were removed for analysis and subsequently replaced.

Demonstration Of The Method.

a. Apparatus. This demonstration of the method for TACOM involved the conversion of a 2-½ ton vehicle and a jeep. A second 2-½ ton vehicle was converted on a trial run basis prior to the demonstration. An M-880 and a 5 ton truck were available for this demonstration but were not converted.

Phase Diagram For The BFS/Polyglycol/2-EH.

a. Apparatus. This diagram was prepared by the titration method, using the fully formulated brake fluids (10).

Infrared Examination Of Hydrogen Bonding Effects.

a. The instrument used for the infrared examination of hydrogen bonding effects in these fluids was a Perkin Elmer 580 using various cells.

Results Of Tests

Air Flushing. The results using three flushing techniques are given in Table 2. With air flushing and vacuum suction (at the bleeder), the residual polyglycol levels are high.

Solvent Addition. This experiment demonstrated the density inversion process. The 2-EH, being a density modifier, reduces the density of the polyglycol to such an extent and in such a manner that upon addition of the silicone, two layers form with the silicone on the bottom. This phenomenon provides the basis for effective polyglycol replacement, the mechanism of which is displacement as opposed to dilution. A critical element in this approach is to achieve complete mixing of the density modifier with the polyglycol fluid especially at the lower part of the wheel cylinders (and calipers).

Wheel Cylinder And Plexiglass Windows. The results of these experiments are outlined in Table 3. The air flushing method for mixing was found to be effective in this system. These experiments established the feasibility of the method and determined that air purging was the method of choice for mixing the density modifier with the polyglycol.

Mock-Up Of A Brake System. The results of these tests are given in Table 4 and a simultaneous flushing technique was found to be feasible in the mock-up.

Caliper Experiments. These experiments demonstrated that a single flush was not sufficient for calipers.

Used Wheel Cylinder. A used wheel cylinder from a 5 ton truck was flushed with 400 ml of 2-EH and emptied into a flask. After the addition of BFS (500 ml), the upper layer contained the particles of sludge broken loose by the flushing. After shaking, a single layer formed and the sludge, to some extent, appeared to have broken up the two phase system when agitated.

Conversion Of An Administrative Vehicle. Table 5 lists the results of the analysis of the fluid from each wheel cylinder and caliper as well as the fluid used. Clearly the method was found to be an effective method for the replacement of polyglycol brake fluid by silicone brake fluid.

Demonstration Of The Method. The results of the analysis of the fluid from the 2-½ ton vehicle used in the dry run of the method are shown in Table 6. Sludge from the walls of the driver side front wheel cylinder did break loose and drained out of the cylinder during disassembly. The observation of a measurable amount of glycol in one of the wheel cylinders is probably due to entrainment of the polyglycol fluid by the sludge, which would render it inaccessible to dissolution by the solvent.

Table 4. Results Of Mock-Up Experiments.

Experiment	Flushes	Residual Alcohol	Residual Polyglycol
Short Lines	3	-	Traces
Long Lines	3	0.96-1.49%	No Trace
Long Lines	2	1.19-6.30%	No Trace
Long Lines	1	1.47-4.29%	Traces

Note: All wheel cylinders were flushed simultaneously.

Table 5. Results Of Conversion Of An Administrative Vehicle.

Wheel	Alcohol Used	Fluid Collected	Residual Polyglycol	*VLT
PR	145 ml	250 ml	None	435°F
DR	145 ml	130 ml	None	453°F
PF	870 ml	285 ml	None	490°F
DF	580 ml	275 ml	None	480°F

Note PR= Passenger rear, etc.,*VLT= Vapor lock temperature.

Table 6. Fluid Analysis From Demonstration With A 2-½ Ton Vehicle.

Wheel Cylinder	Residual 2-EH (%)		Polyglycol
	200 ml	230 ml	
PR1	2.3	0.8	NT
PR2	0.95	0.7	2%
PF	1.6	0.6	T
DF	1.0	0.8	NT
DR1	2.5	0.9	T
DR2	1.7	0.95	T

Note: DF= Driver front, 2-EH column is the volume of BFS passed through the system after conversion, NT= No trace.

Miscellaneous Testing. Table 7 lists a summary of additional testing and evaluation of the use of 2-EH as a flush fluid.

Phase Diagram Determination. The phase diagram was determined using the titration method for the system BFS/Polyglycol/2-EH and Figure 3 shows the diagram obtained. There is a high phase boundary and a low isopycnic line. The BFS and polyglycol used were the fully formulated brake fluids. The working area is above the isopycnic and below the phase boundary. Below the isopycnic, the silicone will be the upper phase. Above the phase boundary a single phase will form (displacement by dilution). There is a slight concavity on the right side of the curve which is assumed to be the result of some additive. The majority of the possible combinations are within the working area and this provides the flexibility needed to develop a method for different vehicles using simple procedures.

Infrared Examination Of Hydrogen Bonding Effects. The -OH stretching bands in dilute solutions were examined. Table 8 lists the bands as well as the spectral shifts due to hydrogen bonding. The spectral shifts for the 2-EH in silicone is smaller than that in glycol which indicates (11) that a weaker hydrogen bond forms in the silicone and may partially explain the preference of the cosolvent for the polyglycol which generates the high phase boundary curve.

Conclusions

A method was developed which is effective at complete polyglycol removal and is illustrated on the right side of Figure 4. The method makes use of an intermediate fluid (2-ethyl hexanol, 2-EH) whose properties are such that a reversal of the phases is induced. Air is used to mix this density modifier with the polyglycol prior to adding the BFS. The method is thus based on the generation of a system which exibits an isopycnic (or twin density) tie line in the phase diagram of the multicomponent binary phase system. While the 2-EH is a solvent for both the BFS and the polyglycol, it greatly prefers the polyglycol in a ternary mixture as is evident from the height of the phase boundary curve (Figure 3). The frequency shifts of the -OH stretching bands indicate a weaker hydrogen bonding of the 2-EH with the silicone as opposed to the polyglycol or the 2-EH itself and may partially explain this preference. In addition, since this procedure does not involve and equilibrium binary phase system, the phase boundary curve may be considered to be the worst case in conducting a conversion using this method.

In addition to having other necessary properties such as a high boiling point, a low freezing point, low flammability, and low moisture absorption, the 2-EH is cost-effective, non-toxic, and is compatible with brake system elastomers in the concentrations generated by this method.

Table 7. Results Of Additional Testing.

Property	2-EH	Parameter In MIL-B-46176 (BFS)	Level 2-EH Tested	Value	Result
Boiling Point	364.6°F	Vapor Lock Temp.	2%	450°F	Pass
Freezing Point	104.8°F	Pass at -65°F	-	-	Pass
Flash Point	185°F	400°F	-	-	Pass
Elastomer Compatability	-	As Stated	2%	-	Pass
Stroking Performance	-	As Stated	3%	-	Pass
Density (g/cm^3)	0.8323	-	-	-	Pass
Moisture Absorption	2.7%	Wet Vapor Lock Temperature After Humidification (0.1% water by Karl Fischer)	5%	360°F	Pass
Toxicity	Slight	-	-	-	Pass
Cost	$3.15/Gal. (bulk rate)	$15-17/Gal.	-	-	Pass

Note: The minimum flash point for VV-B-680 fluids is 179.6°F.

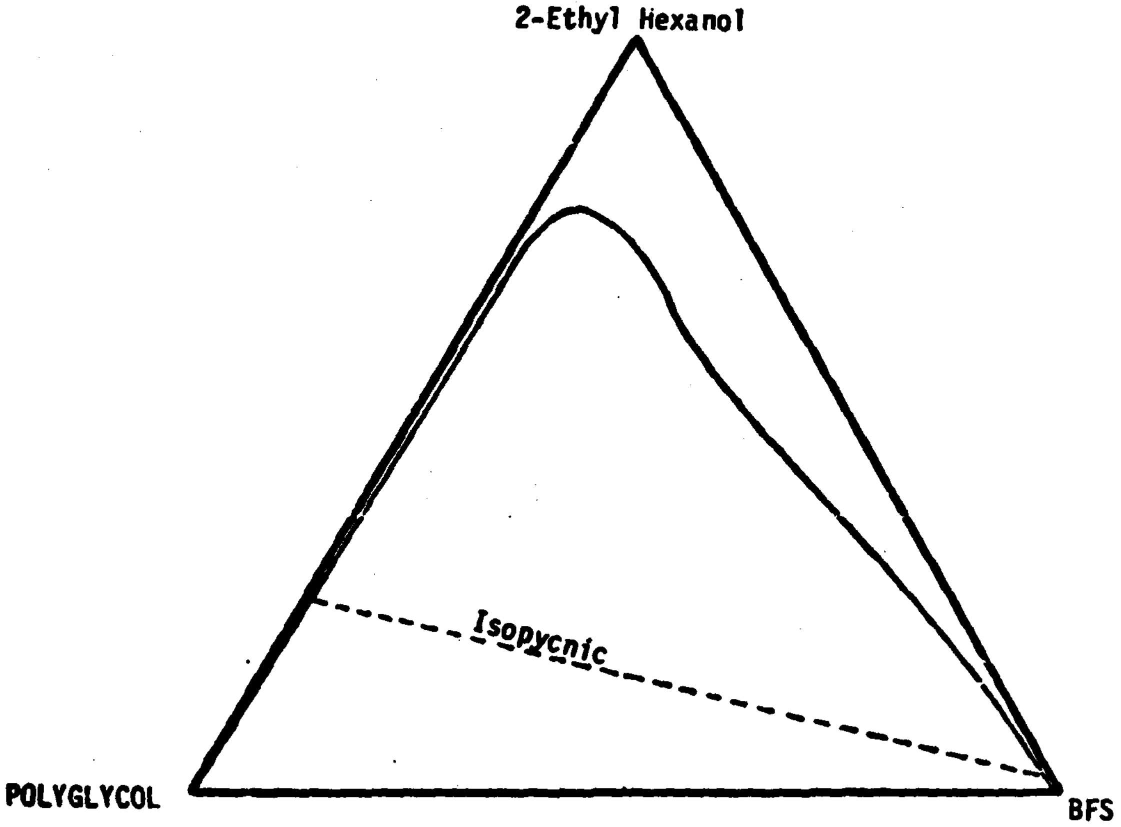

2-Ethyl Hexanol
Isopycnic
POLYGLYCOL
BFS

Figure 3. Phase Diagram For The BFS/Polyglycol/2-EH System. The fully formulated fluids were used. The slight concavity on the right side of the diagram is assumed to be due to additives. Note the high phase boundary and the low isopycnic (dashed line).

Table 8. Hydroxyl Stretching Frequencies For Alcohols.

Concentration (%)	Monomer	Dimer	Polymer	
2-EH/CCl_4				
2	3639	3499	3359	280
1	3639	3523	-	126
0.5	3639	3510	-	129
neat	3640	-	-	300
2-EH/BFS				
2	-	-	3400	239
10	-	-	3440	199
Cyclopentanol/CCl_4				
2	3624	3490	3357	267
1	3624	3495	3365	259
0.5	3624	3505	3365	259
neat	-	-	3340	
Cyclopentanol/BFS				
Saturated Solution	-	-	3340	284
n-Butanol/CCl_4				
2	3636	3494	3344	292
1	3636	3503	3360	276
0.5	3638	3511	-	127
neat	-	-	3340	296

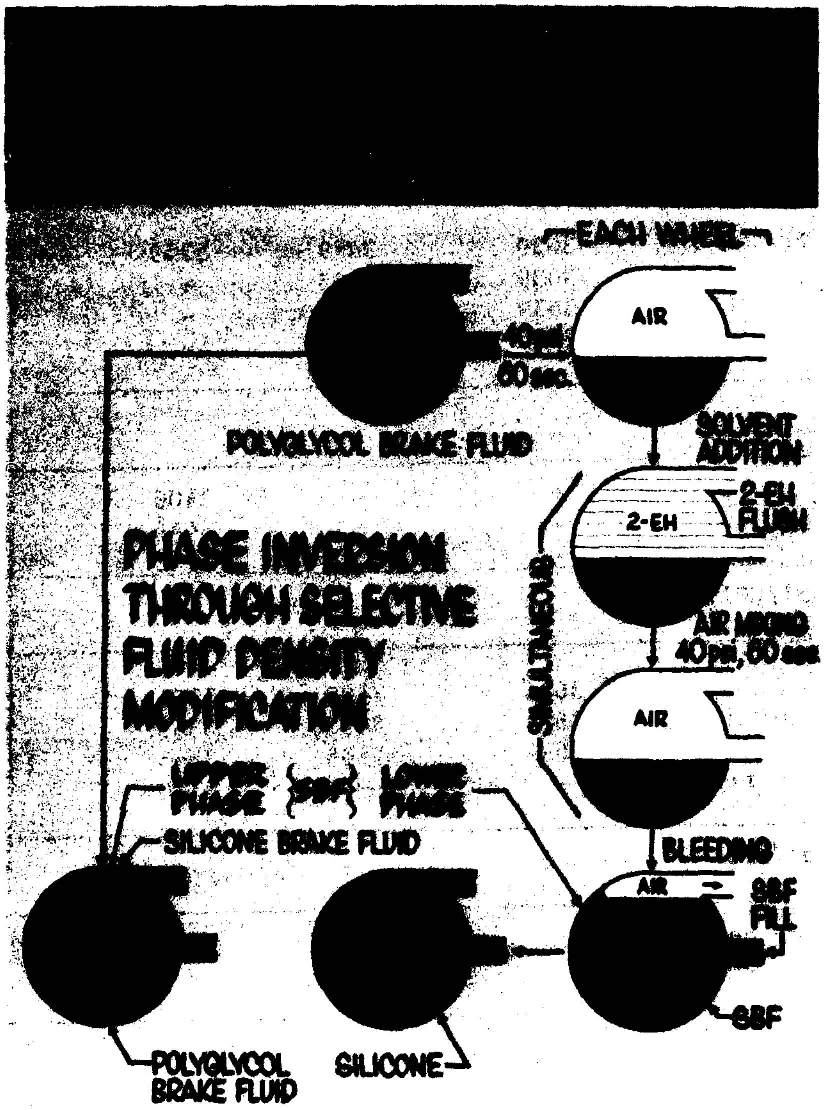

EACH WHEEL
AIR
40psi
60sec.
POLYGLYCOL BRAKE FLUID
SOLVENT ADDITION
2-EH
2-EH FLUSH
PHASE INVERSION THROUGH SELECTIVE FLUID DENSITY MODIFICATION
SIMULTANEOUS
AIR MIXING
40psi, 60sec.
AIR
UPPER PHASE
SBF
LOWER PHASE
SILICONE BRAKE FLUID
BLEEDING
AIR
SBF FILL
SBF
POLYGLYCOL BRAKE FLUID
SILICONE

Figure 4. Comparison of Flush/Fill and Solvent Assisted Flush/Fill Procedures For The Conversion of Tank-Automotive Equipment to Brake Fluid, Silicone (BFS).

The USMC will be using the method (which is being patented) to convert their shipboard vehicles on board the Near Term Prepositioned Fleet at Diego Garcia to BFS. An extension of the method covering mineral fluid based systems is being considered by industry. The conversion to BFS is expected to save the US Army $21M over the next 10 years. This method could not only recapture the all climate utility and high vapor lock temperature of BFS but could substantially increase the cost-effectiveness of the conversion.

References

1. Taken in part from the Ph.D. thesis of C.C.C., The American University, Washington, D.C., May, 1982.

2. Federal Specification VV-B-680, Brake Fluid, Automotive, July 20, 1972; Military Specification MIL-H-13910, Hydraulic Fluid, Polar Type, Automotive, All Weather, February 3, 1967; Military Specification MIL-P-46046, Preservative Fluid, Automotive Brake System and Components, August 26, 1964.
3. Conley, J.H., and Jamison, R.G., Army Experience With Silicone Brake Fluid, S.A.E. Technical Paper Series #780660, 1978.

4. Military Specification MIL-B-46176, Brake Fluid, Silicone, Automotive, All Weather, Operational and Preservative, March 27, 1978.

5. Conley, J.H., and Jamison, R.G., Silicone Brake Fluid: One Year Field Test, MERADCOM Report No. 2132, February 1975.

6. TB43-0002-87, DA Technical Bulletin, Brake Fluid, Silicone (BFS) Conversion Procedures for Tank-Automotive Equipment, May 1981.

7. MERADCOM Letter Report, Evaluation of Flush/Fill Procedures for Silicone Brake Fluid Retrofit, DRDME-GL, October 6, 1980.

8. Francis, A.W., Ind. Eng. Chem., 45, 2789, 1953.

9. Smith, A.L., ed., Analysis of Silicones, Vol. 41, Cemical Analysis Series of Monographs, Wiley, New York, 1974.

10. Francis, A.W., Liquid-Liquid Equilibriums, Vol 3, Interscience Library of Chemical Engineering and Processing, Interscience Publishers, New York, 1963.

11. Pimentel, G.C., and McClellan, A.L., The Hydrogen Bond, Freeman, San Francisco, 1960.

A MORE RATIONAL APPROACH TO THE STRESS ANALYSIS OF PROJECTILES

*SHIH C. CHU, PhD
JAMES STEINER, MR.
TECHNOLOGY BR, ARMAMENT DIV, FC&SCWSL
US ARMY ARMAMENT RESEARCH AND DEVELOPMENT COMMAND
DOVER, NEW JERSEY 07801

INTRODUCTION

Projectiles are traditionally analyzed and designed by using theory of elasticity [1,2,3]. In this approach, the entire projectile body is considered to be loaded within the elastic region. However, under actual firing conditions, the equivalent stress in some region of the projectile is much greater than the yield strength of the projectile material. Hence, plastic flow will be encountered in the projectile body. The present trend in stress analysis of weapon and ammunition components is to take into account, in a more rigorous fashion, the complex phenomena of plastic flow. This is due to the necessity of designing for maximum stress to achieve an optimal design. In this investigation, a more rigorous nonlinear technique is developed in order to predict the inelastic deformation and stress distributions of a projectile subjected to actual firing conditions. Both nonlinear material response and geometric nonlinearity have been taken into consideration. Nonlinearity of material properties has been taken into account by use of theories of plasticity. Geometric nonlinearity has been considered by use of the finite element approach. A complete inelastic stress analysis of a 30 mm XM TP projectile has been conducted. The intent of this investigation is to identify potential design flaws and critical regions of the projectile undet actual firing environment. A load history was defined which subjected the projectile to the loads at various points on the barrel travel curve. A peak pressure of 60,000 psi, peak acceleration of 4.8×10^7 in./sec^2 and a maximum spin of 6,355 rad/sec were considered.

CONSTITUTIVE EQUATIONS

The primary objective of this investigation is to develop a more rational, nonlinear elastoplastic method for analyzing and designing a projectile to achieve in-bore structural integrity. The incremental stress-strain relations associated with the Von Mises yield criterion obeying the kinematic-hardening law for work-hardening materials will be used. The kinematic hardening law in the incremental theory of plasticity was originally proposed by Prager [4], for the deformation of isothermal solids. Later, Prager [5] extended these formulations to the nonisothermal condition for rigid work-hardening solids. Recently, Chu [6,7], expanded Prager's work for solids of more general deformation state. A brief outline of the constitutive equations used in the analysis is now given.

The total increment strain tensor, $d\varepsilon_{ij}$ is assumed to be the sum of elastic part, $d\varepsilon^e_{ij}$, plastic part, $d\varepsilon^P_{ij}$, and the part of thermal strain, $d\varepsilon^T_{ij}$, i.e.

$$d\varepsilon_{ij} = d\varepsilon^e_{ij} + d\varepsilon^P_{ij} + d\varepsilon^T_{ij} \tag{1}$$

The elastic strain components $d\varepsilon^e_{ij}$ are related to the incremental stress components, $d\sigma_{ij}$, by

$$d\sigma_{ij} = D_{ijkl}d\varepsilon^e_{kl} = D_{ijkl}(d\varepsilon_{ij} - d\varepsilon^P_{kl} - \beta dT\delta_{kl}) \tag{2}$$

in which β is the thermal expansion coefficient, dT, is the change in temperature, and δ_{kl} is Kronecker delta. For isotropic material, the 4th rank material tensor, D_{ijkl} is defined by

$$D_{ijkl} = \frac{E}{1+\nu}\ \delta_{ik}\delta_{jk} + \frac{\nu E}{(1+\nu)(1-2\nu)}\ \delta_{ij}\delta_{kl} \tag{3}$$

where E is Young's modulus and ν is Poission's ratio of material.

On the basis of the Von Mises yield criterion with temperature-dependent yield strength of a material, the yield surface can be represented as:

$$f = \frac{1}{2}\, S_{ij}S_{ij} - \kappa^2(T) = 0 \tag{4}$$

where

$$S_{ij} = (\sigma_{ij} - \alpha_{ij}) - \frac{1}{3}\,(\sigma_{kk} - \alpha_{kk})\delta_{ij}; \qquad S_{mm} = 0 \tag{5}$$

α_{ij} is a tensor representing the total translation of the center of the initial yield surface, and κ is related to the uniaxial yield stress $\kappa = \sigma_y(T)/\sqrt{3}$.

In addition to the yield condition, a constitutive relation between plastic strain increments, stress, and stress increments is required to describe the inelastic behavior of a material. The constitutive relation (flow rule) used in this investigation is based on Drucker's postulate for work-hardening material [8]. The flow rule is given as:

$$d\epsilon_{ij}^{P} = d\lambda\frac{\partial f}{\partial\sigma_{ij}} \tag{6}$$

where $d\lambda$ is a positive scalar quantity. On the basis of Drucker's statement, this plastic strain increment tensor must lie on the outward normal to the yield surface at the instantaneous stress state.

Based upon Prager's kinematic-hardening rule [4,9] with Ziegler's modification [10], the increment of translation of the center of the yield surface is assumed to be directed along the radius vector connecting the center of the yielding surface to the instantaneous stress state, i.e.,

$$d\alpha_{ij} = (\sigma_{ij}-\alpha_{ij})d\mu; \; d\mu>0 \tag{7}$$

where $d\mu$ can be determined provided the stress point remains on the tranlated yield surface during plastic flow, i.e.,

$$d\mu = \frac{\frac{\partial f}{\partial\sigma_{kl}}\, d\sigma_{kl} + \frac{\partial f}{\partial T}dT}{(\sigma_{mn}-\alpha_{mn})\;\frac{\partial f}{\partial\sigma_{mn}}} \tag{8}$$

During plastic loading, the consistency condition requires that

$$df = \frac{\partial f}{\partial\sigma_{ij}}(d\sigma_{ij}-Cd\epsilon_{ij}^{P}) + \frac{\partial f}{\partial T}dT = 0 \tag{9}$$

The plastic strain vector $Cd\epsilon_{ij}^{P}$ is considered as the projection of $d\sigma_{ij}$ (and thus of $d\alpha_{ij}$) on the exterior normal to the yield surface, where C is a material constant. Hence, for small incremental of stress and strain, one can readily find that

$$d\lambda = \frac{1}{C} \frac{\frac{\partial f}{\partial \sigma_{ij}} d\sigma_{ij} + \frac{\partial f}{\partial T} dT}{(\frac{\partial f}{\partial \sigma_{mn}})(\frac{\partial f}{\partial \sigma_{mn}})} \qquad (10)$$

Therefore, the flow rule becauses

$$d\varepsilon_{ij}^{P} = \frac{\frac{\partial f}{\partial \sigma_{ij}}}{C(\frac{\partial f}{\partial \sigma_{mn}})(\frac{\partial f}{\partial \sigma_{mn}})} (\frac{\partial f}{\partial \sigma_{kl}} d\sigma_{kl} + \frac{\partial f}{\partial T} dT) \qquad (11)$$

If axisymmetric deformation is considered, with reference to the cylindrical coordinates (r, Θ, z) the state is defined by the nonvanishing stress components $\{d\sigma\}^T = < d\sigma_r, d\sigma_\Theta, d\sigma_z, d\tau_{rz} >$ and strain components $\{d\varepsilon\}^T = < d\varepsilon_r, d\varepsilon_\Theta, d\varepsilon_z, d\gamma_{rz} >$. Then, the incremental stress-strain relations are found in the following matrix form:

$$\{d\sigma\} = [\overline{D}]\ \{d\varepsilon\} - \beta dT\ \{B\} \qquad (12)$$

in which,

$$[\overline{D}] = \frac{\lambda}{\nu} \begin{bmatrix} 1-\nu & \nu & \nu & 0 \\ \nu & 1-\nu & \nu & 0 \\ \nu & \nu & 1-\nu & 0 \\ 0 & 0 & 0 & \frac{1-2\nu}{2} \end{bmatrix} - \eta_1 \cdot \begin{bmatrix} S_r^2 & S_r S_\theta & S_r S_z & S_r \tau_{rz} \\ S_\theta S_r & S^2{}_\theta & S_\theta S_z & S_\theta \tau_{rz} \\ S_z S_r & S_z\ S_\theta & S_z^2 & S_z \tau_{rz} \\ \tau_{rz} S_r & \tau_{rz} S_\theta & \tau_{rz} S_z & \tau_{rz}^{\ 2} \end{bmatrix} \qquad (13)$$

$$\{B\} = \frac{E}{1-2\nu} \begin{Bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{Bmatrix} - \frac{\eta_2}{\beta} \cdot \frac{\partial \kappa}{\partial T} \begin{Bmatrix} S_r \\ S_\theta \\ S_z \\ \tau_{rz} \end{Bmatrix} \qquad (14)$$

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)} \qquad (15)$$

$$\eta_1 = \frac{1}{g} \left(\frac{E}{1+\nu}\right)^2 \tag{16}$$

$$\eta_2 = \frac{2\kappa}{g} \cdot \frac{E}{1+\nu} \quad \text{and,} \tag{17}$$

$$g = 2\kappa^2 \cdot \left(\frac{E}{1+\nu} + C\right) \tag{18}$$

METHOD OF SOLUTION

The projectile and all loading acting on the projectile are considered as axisymmetric. A cylindrical coordinate system (r, θ, z) is used in this analysis. By the assumption of axial symmetry, all variables are independent of angle θ, consequently, all derivatives, with respect to θ vanish. The displacement u_θ, and the shear stresses $\tau_{r\theta}$ and $\tau_{\theta z}$ vanish.

Due to the complexity of the geometry of a projectile and nonlinear material behavior, the finite element method [11,12] was used to conduct the stress analysis. The analytical approach used in this investigation is the incremental loading technique, wherein at each step of loading a new stiffness matrix is formulated, in terms of the finite element model, and solved for incremental deformations, stresses, and strains.

To perform a finite element stress analysis, the cross section of a projectile is divided into a large number of small triangular and quadrilateral elements as shown in Figure 1. Only a few nodal points on the boundary need be specified, the remaining nodal points are obtained from an automatic mesh generating computer program. The grid was partitioned to place a finer grid at those areas of the body that are expected to undergo large stress.

PROJECTILE CONFIGURATION AND MATERIAL PROPERTIES

The outline of a 30 mm XM TP projectile is shown in Figure 2. The projectile is made of three distinct parts with three different materials as shown in Figure 2 (the projectile body, nose cap, and rotating band are made of 1018 steel, aluminum, and gilding metal, respectively). The material properties are given below:

Material	Poisson's Ratio	Young's Modulus (psi)	Yielding Stress (psi)	Ultimate Stress (psi)	Density lb/in.^{3}
Steel	.29	29×10^6	88,500	90,000	.280
Aluminum	.33	10×10^6	77,000	80,000	.101
Gilding Metal	.33	17×10^6	40,000	48,000	.317

The stress-strain relations of those materials are shown in Figure 3.

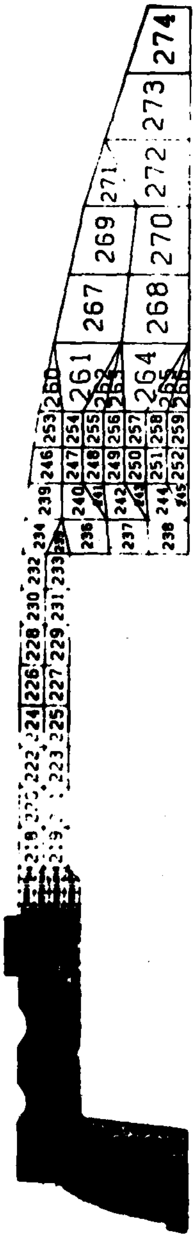

274
273
272
271
270
269
268
267
264
261
260
259
258
257
256
255
254
253
252
251
250
249
248
247
246
244
242
240
239
238
237
236
234
233
232
231
230
229
228
227
226
225
224
223
222

Figure 1. Finite Element Presentation of a Projectile

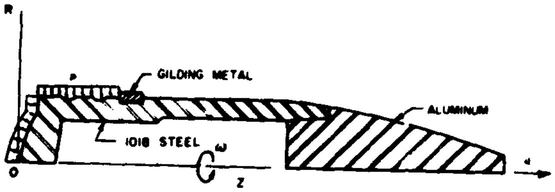

R
P
GILDING METAL
ALUMINUM
1018 STEEL
ω
O
Z

Figure 2. Projectile Configuration

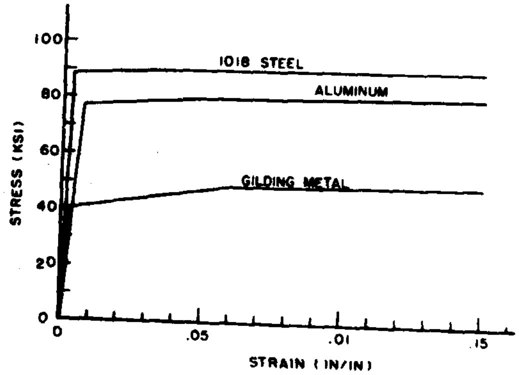


Figure 3. Stress-Strain Relation for Projectile Materials

LOADING CONDITIONS

The loads considered in this investigation simulating the environment in the gun barrel during firing. The loading consisted of four types: (1) propellant gas pressure, (2) setback force due to acceleration, (3) rotational velocity due to the spin of projectile, and (4) forced displacement of the rotating band. The maximum value of these four types of loadings are given by, P_{max} = 60,000 psi, a_{max} = 4.8×10^7 in./sec^2, ω_{max} = 6,335 rad/sec, and Δ_{max} = -0.008 inches for gas pressure, acceleration, spin rate, and radial displacement, respectively. The frictional shear forces between the band and the barrel were neglected in this investigation. The loads were applied in incremental fashion with the relative magnitudes at each load point simulated the physical interdependence of the loads at different times during the interior ballistic cycle. The detailed loading history is given in Figure 4.

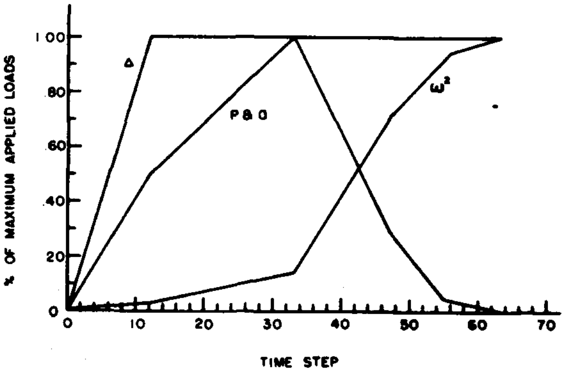


Figure 4. Loading History

COMPUTATIONAL RESULTS

For each incremental loading, the stresses and strains of each element and the displacement of each nodal point were calculated. The plastic flow has been initiated within element 186 of the gilding metal at the end of the first step of loading. Based upon the Von Mises yield criteria, the equivalent stress for each element was calculated. At the end of the first step of loading, the maximum equivalent stress (σ_e) in the projectile body is 69,900 psi which is lower than yielding strength of 1018 steel (element 174), and in the gilding metal portion is 40,000 psi (element 186), which is equal to the yield strength of the gilding material.

The computation results presented in this section are based upon the following three major loading steps (shown in Figure 4):

1) At the end of the 12th step ($\Delta=1.0$, $p=0.5$, $a=0.5$, and $\omega^2=0.036$) while the forced displacement reeaches the maximum in the rotating band.

2) At the end of the 33rd step, ($\Delta=1.0$, $p=1.0$, $a=1.0$, and $\omega^2=0.14$) while both propellant gas pressure and setback force reach the peak.

3) At the end of the 63rd step ($\Delta=1.0$, $p=0$, $a=0$, and $\omega^2 = 1.0$), while the spin rate reaches the maximum at the muzzle end of the tube.

The plastic range in the projectile under the above three loading conditions is shown in Figure 5, 6, and 7, as indicated by the shaded area.

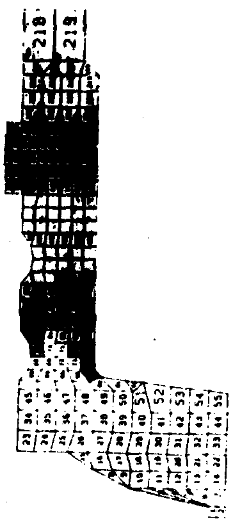


Figure 5. Plastic Zone in the Projectile at Loading Condition: $\Delta=1.00$, $p=0.50$, $a=0.50$ and $\omega^2=.036$

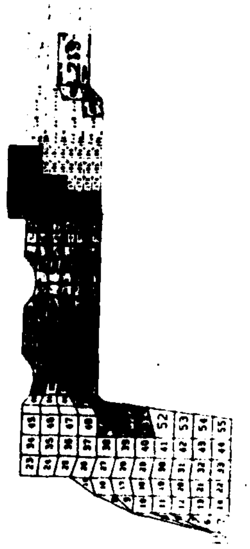


Figure 6. Plastic Zone in the Projectile at Loading Condition: $\Delta=1.00$, $p=1.00$, $a=1.00$ and $\omega^2=0.14$

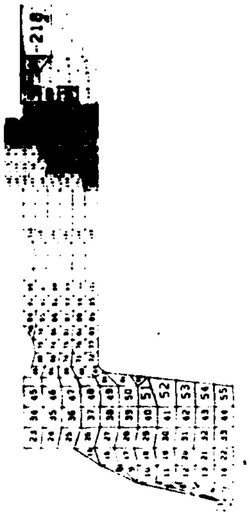
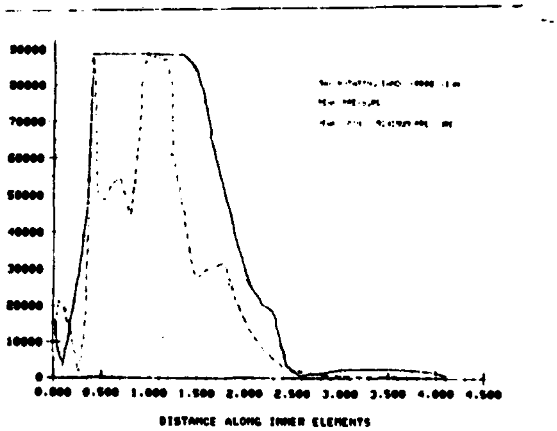


Figure 7. Plastic Zone in the Projectile at Loading Condition: Δ=1.00, p=0.0, a=0.0 and ω^2=1.00

The maximum plastic strains encountered in the gilding metal (element 186) is 0.027 in./in., and in the projectile is 0.025 in./in. (element 163) at the end of the 12th loading step, while the maximum radial displacement took place in the rotating band. The maximum plastic strains were increased to 0.0294 in./in. (element 186) in the gilding metal and to 0.0254 in./in. (element 163) in the projectile body, while both gas pressure and setback force reached the peak. When projectile reached the muzzle end the maximum plastic strains are 0.0294 in./in. and 0.0258 in./in. encountered in the elements 186 and 163, respectively. During the entire loading process the change of maximum plastic strain is small. The location of plastic zones under different loading conditions is important to the projectile design.

The equivalent stress along the length of the inner surface elements of the projectile is shown in Figure 8. The maximum equivalent stress reaches 88,800 psi, which is greater than the yielding strength (σ_y = 88,500 psi) and smaller than the ultimate strength (σ_u = 90,000 psi) of the projectile material. If the investigation was based on elastic analysis, the maximum equivalent stress could reach a value which would be much higher than the ultimate strength of the projectile material. This is the reason that our inelastic analysis technique be developed, since a more rigorous stress analysis can only be obtained by using nonlinear theories of plasticity as developed in this investigation.



See Next Page for Enlarged Scale

Figure 8. Equivalent Stress in the Inner Elements of the Projectile.

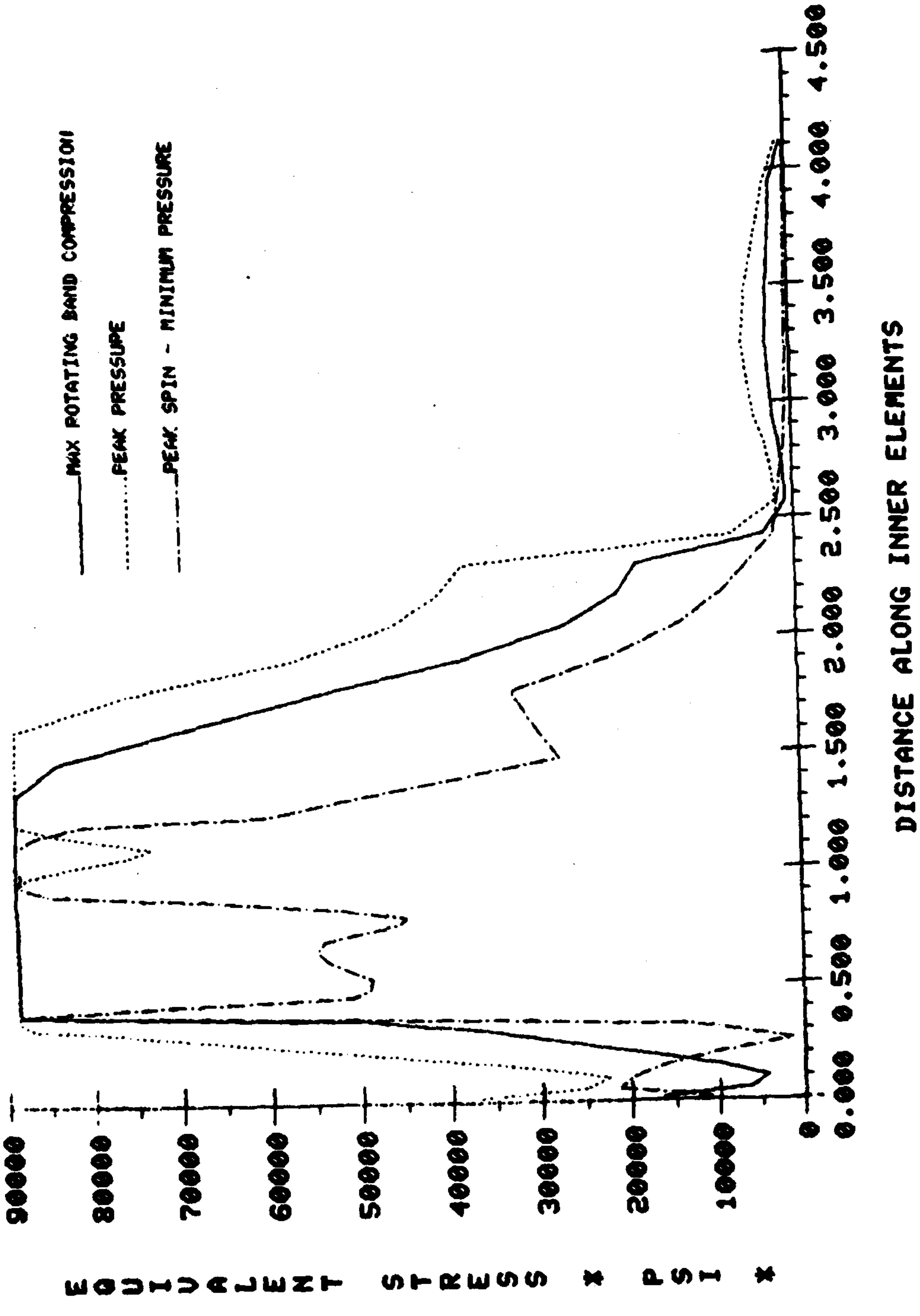


Figure 8 . Equivalent Stress in the Inner Elements of the Projectile.

The equivalent stress distribution in the projectile of high stressed region at peak pressure and setback is shown in Figure 9. It is clearly indicated that the critical region is located from the rotating band to the inside wall of projectile base.

The deformed grids under different loading conditions are shown in Figures 10, 11, and 12. The deformations shown in these Figures are the deformations enlarged 10 times in magnitude.

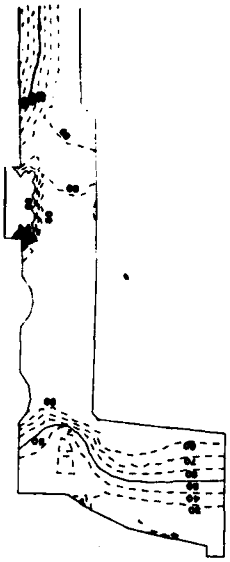


Figure 9. Equivalent Stress (KSI) Distribution in the Critical Region of the Projectile at Loading Condition: $\Delta=1.0$, $p=1.0$, $a=1.0$ and $\omega^2=0.14$

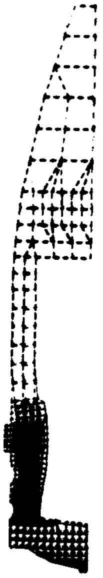


Figure 10. Deformed Grid Mesh Under Loading Condition: $\Delta=1.0$, $p=0.5$, $a=0.5$ and $\omega^2=0.036$

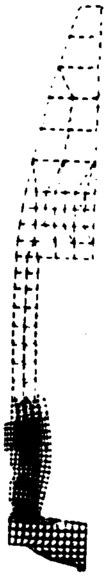


Figure 11. Deformed Grid Mesh Under Loading Condition: $\Delta=1.0$, $p=1.0$, $a=1.0$ and $\omega^2=0.14$

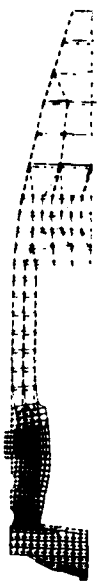


Figure 12. Deformed Grid Mesh Under Loading Condition: $\Delta=1.0$, $p=0$, $a=0$ and $\omega^2=1.0$

CONCLUSIONS

On the basis of Drucker's flow rule of plasticity, the Von Mises flow criterion, and the strain-hardening and compressibility properties of a material, a more rational rigorous nonlinear elastoplastic analytic method has been develped for analyzing and designing a projectile subjected to actual complicated firing conditions. In constrast to the traditional designing and analyzing techniques (theories of elasticity), the complex phenomenal plastic flow in the projectile has been taken into account. This is due to the necessity of designing for maximum stress to achieve an optimal design for reducing component weight. Both nonlinear material response and geometric nonlinearity have been taken into consideration. Nonlinearity of material properties has taken into account by use theories of platicity. Geometric nonlinearly has been considered by use the finite element technique. An incremental loading procedure has been used to consider the actual firing environment of a gun. The complete loading history which includes propellant gas pressure, setback force, and spin rate of a projectile was defined as function of time.

The critical region is located in the region between rotating band and projectile base. In this region the equivalent stress in general is above the yield strength of material, however, it is below ultimate strength of the material. The plastic strain has taken place in the reigon, however, maximum equivalent plastic-strain at inner surface is below 0.03 in./in., which is relatively small.

Based upon our investigation the plastic zone in the projectile has been identified and accurately located for each incremental of loading which simulating the actual fire environment in a gun. The developed technique will provide a more rigorous stress analysis tool for projectile design. The potential design flaws and critical regions of a projectile can be identified before the projectile is being made and tested.

REFERENCES

1. DePhillipo, T.E., and Booth, A.W., "Mathematical Model for Determining Stresses in Projectile Bodies," Report R-1939, Frankford Arsenal, Oct 1969.

2. Mechanical Engineering Department, New York University, "Principal and Combined Sresses of the Shell, HE 155 mm M197," Picatinny Arsenal Report U38115.

3. Elder, A.S., Burns, B.P., and Hurban, J.M., "Stress Analysis of 175 mm Projectile, HE M437, BRL Memorandum Report 2113.

4. Prager, W., "The Theory of Plasticity: A Survey of Recent Achievement (Hames Clayton Lecture)," Proc. Inst. Mech. Eng. 69 (1955).

5. Prager, W., "Nonisothermal Plastic Deformation," Proc. Koni, Ned. Akad., Van Weton, Series B, 61, NO (3), 1958.

6. Chu, S.C., "Nonisothermal Elastoplastic Deformation of Work-Hardening Solids," Paper presented at Joint ASME/ASCE Mechanics Conference, University of Colorado, 22-24 June 1981.

7. Chu, S.C., "An Incremental Appraoch to Nonisothermal Elastic-Plastic Deformation." Proc. 27th Conference of Army Mathematicians, 1981.

8. Drucker, D.C., "A More Fundamental Approach to Plastic Stress-Strain Relations," Proc. 1st U.S. National Congress Applied Mechanics, New York, 1952.

9. Prager, W., "A New Method of Analyzing Stresses and Strains in Work-Hardening Plastic Solids," Journal App. Mech. Vol. 23, 1956.

10. Ziegler, H., "A Modification of Prager's Hardening Rule," Quarterly of Appl. Math Vol. 17, 1959.

11. Zienkiewicz, O.C., "The Finite Element Method in Engineering Science." McGraw-Hill, 1971.

12. Marcal, P.V. and King, I.P., "Elastic-Plastic Analysis of Two-Dimensional Stress Systems by the Finite Element Method," Inter. J. Mech., Vol. 9, No. 3, 1967.

MUNITION EXPENDITURE MODEL VERIFICATION: KWIK PHASE I (U)

*STEPHEN L. COHN
RICARDO PENA
US ARMY ATMOSPHERIC SCIENCES LABORATORY
WHITE SANDS MISSILE RANGE, NEW MEXICO 88002

The threat imposed by the Soviet bloc tank forces requires that ground level obscuration, for both offensive and defensive planning, must receive the most thorough research and developmental efforts. The possibility of reducing costs for munition expenditures further dictates pursuit of more efficient mechanisms for obtaining obscuration objectives.

A large data base from previous experiments exists in the literature, covering chemically generated military smokes. These data have been used to verify and/or evaluate several different approaches to atmospheric diffusion, including the Gaussian formulae. These previous tests have verified the predictability of relatively long average downwind concentrations of some diffusing materials in the atmosphere. However, a deficiency exists in the case of military smokes for which the actual obscuration has not been reliably predicted or verified, especially over short time intervals. This deficiency has made it impossible to evaluate KWIK, a munition expenditure model, in all categories of performance without obtaining additional data.

In order to verify the munition expenditure predictions of the KWIK model, an evaluation plan consisting of three phases was devised by Atmospheric Sciences Laboratory (ASL). This paper deals only with Phase I, an effectiveness evaluation test for visible wavelengths.

1. DESCRIPTION OF MODEL

The KWIK (a mnemonic derived from crosswind integrated concentration) model consists of a blending of meteorological and site parameters, atmospheric optics, and turbulent diffusion theory. Each of these are briefly discussed below.

Meteorological data requirements for the KWIK algorithm are based upon observations that would be available on a modern battlefield (i.e., hourly airway type data obtained from the United States Air Force (USAF) Air Weather Service via the USAF Global Weather Central, or information furnished by the US Army Field Artillery Meteorological Sections).

Observational requirements for the microscale diffusion, atmospheric optics, ambient stability, and wind direction effects upon the obscuring screen were investigated, with the determination being that eight standard meteorological parameters and one terrain characterization index would be sufficient for the KWIK algorithm. The eight meteorological data inputs consist of

a. ceiling height in feet
b. cloud cover in percent
c. visibility in miles
d. precipitation, yes or no
e. temperature in degrees Fahrenheit
f. dew point temperature in degrees Fahrenheit
g. wind direction in degrees (meteorological convention)
h. windspeed in knots

The terrain index is the average height, in centimeters, of the surface roughness elements, such as trees, bushes, grasses, or buildings.

The stability category scheme used is a composite version developed from the published results of Pasquill (1), Turner (2), and Smith (3). The composite approach uses Turner's radiation index, ceiling, and cloud modifications to the index, and Smith's windspeeds associated with each Pasquill category. Other inputs related to the calculation of insolation for the determination of the atmospheric stability category are

a. latitude in degrees
b. direction from equator (north or south)
c. longitude in degrees
d. direction from Greenwich (east or west)
e. altitude above MSL in kilometers
f. Julian date in three digits
g. Greenwich civil time in hours

KWIK contains an optics section that is adapted from an approach to atmospheric transmission by Downs (4). The model assumes that 5 percent (5), (6), or less transmittance will deny acquisition of a target in the visible wavelengths. The transmittance of light at various wavelengths through a path is determined by calculating the attenuation due to both absorption by water vapor and scattering by natural atmospheric aerosols. The concentration of smoke necessary to obscure a target through a given line of sight is then calculated, along with the number of munitions necessary to deliver the required amount of obscurant.

For a continuous smoke source, such as the hexachloroethane (HC) used in the KWIK Phase I trials, the smoke is assumed to have a Gaussian distribution and to diffuse independently in three coordinate directions (X, Y, Z). The crosswind integrated concentration (CWIC) equation used is based on the Gaussian distribution function described by Pasquill (1) and Gifford (7) and modified by Umstead et al (8).

2. DESCRIPTION OF TRIALS

Thirty trials were conducted at Dugway Proving Ground (DPG) during July and September of 1980. Groups of three M1 and one M2 HC smoke canisters were statically detonated to simulate each 155-mm M116BE projectile needed during a given screen.

Test Objectives. The objectives of the KWIK Phase I evaluation test were

a. To provide an evaluation of the KWIK smoke model by correlating model predictions of obscuration effectiveness with empirical (observer) data.

b. To collect meteorological, photographic, and observer data in order to characterize the meteorological, environmental, and smoke plume behavior for each trial.

c. To compare and evaluate smoke munition expenditure calculations of the KWIK model from successful screens with those obtained by the current method used by the field army (9), (10), (11).

d. To compare meteorological data from a distant (10 km) source and evaluate its effect on the munition expenditures calculated by the KWIK model.

Data Requirements. The main data requirements consisted of meteorological, photographic, and visual observational data.

Meteorological data were measured at the test site (horizontal grid) from three different towers (figure 1). Two 10-m towers were located at the southeast and northwest ends of the grid, respectively, and a 32-m tower was located on the northeast side of the grid next to the observation post. Upper atmospheric data were collected at the Ditto

Meteorological Station, located about 10 km east of the test area, at the DPG Ditto Technical Center.

The photographic coverage provided during the trials consisted of three 16-mm color motion picture cameras located as shown in figure 1. Color still photographs were taken every 30 s during each trial with a camera located near the area of the observation post (figure 1).

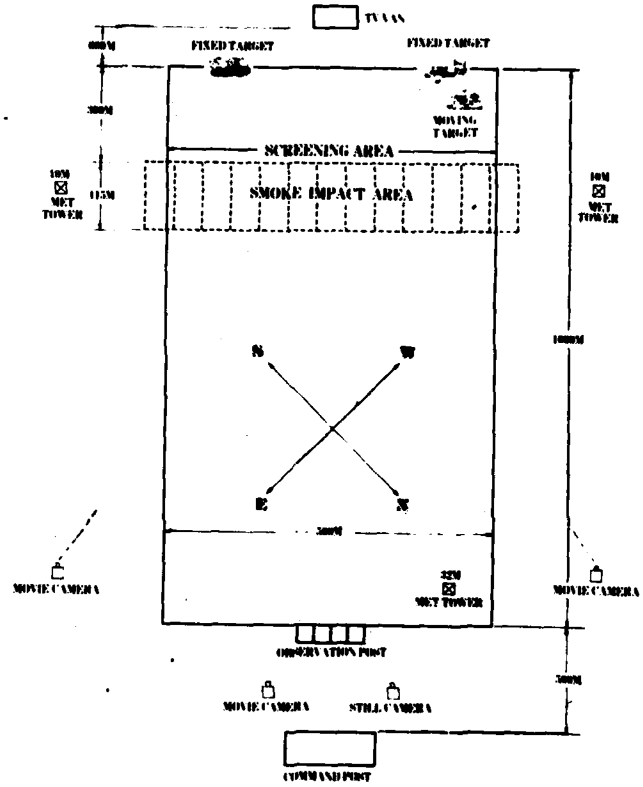

TV VAN
400M
FIXED TARGET
FIXED TARGET
300M
MOVING TARGET
SCREENING AREA
10M MET TOWER
115M
SMOKE IMPACT AREA
10M MET TOWER
1000M
N
W
E
S
300M
MOVIE CAMERA
32M MET TOWER
MOVIE CAMERA
OBSERVATION POST
300M
MOVIE CAMERA
STILL CAMERA
COMMAND POST

Figure 1. Test Grid for KWIK smoke tests, Phase I, DPG, Utah.

Visible smoke obscuration assessments were made from the observation post. Each observer (with binoculars) was situated in a booth and had an unrestricted view of the target area (figure 1). Each of three observers was assigned one of the three targets, with a fourth observer assigned all three targets. The first three observers activated a recording device when their assigned target was visible. The fourth observer activated a recording device when one or more targets were visible.

Smoke Impact Area. This area included the screening area plus 30 m to the southeast and to the northwest, for a total of 560 m (figure 1). Each row consisted of three lines (a, b, and c), on which the required HC smoke canisters were placed. The munitions along each selected "a" line were ignited simultaneously, while the ignition of the "b" and "c" lines were delayed by 2-min time intervals. This arrangement was used to simulate volleys of dynamically fired M116 155-mm HC rounds.

KWIK Calculations. An HP85 desktop calculator located at the command post was used to perform the KWIK smoke model munition expenditure calculations. Using the meteorological and site data from the test grid prior to each trial as inputs, the model produced the outputs that were used for the appropriate trials on a real-time basis.

3. EVALUATION OF DATA

A percentage of "effective screening" was calculated by comparing the total time the observers viewed their targets with the maximum time any target was scheduled to be visible. The evaluation of the data was performed in three parts: the Smoke Screening Assessment, the Munition Expenditure Assessment, and the High Wind (>8 m/s) Smoke Screening Assessment. The Smoke Screening Assessment was based on target obscuration data, with the High Wind Smoke Screening Assessment data coming from trials 23 through 30. One Munition Expenditure Assessment compared the KWIK munition expenditures with those obtained using the method in the current FM (FM 6-40-5) (9), (10). Another Munition Expenditure Assessment compared munition expenditure results using meteorological data collected at two sites: the DPG Horizontal Test Grid and the Ditto Meteorological Station. For the purpose of evaluation, all of the test data were grouped according to windspeed as follows: 2.0 to 3.5 m/s, 3.6 to 7.5 m/s, and >8 m/s.

Smoke Screening Assessment. The target observer assessment data were plotted as a function of time. These data were analyzed for all trials, except trials 1 through 4, which lacked observer data. Photographic data from each trial were used to verify the target observer assessment data.

Low windspeed screens (2.0 to 3.5 m/s) were successful in four out of six cases, with an average effective screening during 88.3 percent of the total target time, as shown in table 1. Trials 16 and 18 were unsuccessful because of low windspeeds (<2 m/s), which contributed to excessive plume rise, preventing formation of a screen. Significantly, the FM method's predictions for trials 16 and 18 were identical to those produced by KWIK, indicating that both models were unable to predict an effective screen.

TABLE 1. SMOKE SCREENING ASSESMENT, WINDSPEED: 2.0 - 3.5 M/S

Trial No.	Target	Initial Time To Screen Target (min: sec)	Approx Build Up (min: sec)	Target Schedule (min: sec)	Time Target Observed (min: sec)	Effective Screening (%)	Mean Effective Screening (%)
11	Tank (T)	*		4:30	0:00	100	
	Jeep (J)	*		2:30	0:00	100	
	Moving (M)	*		5:00	0:45	85	
	All (A)	*	1:00	5:00	1:42	66	88
14	T	*		2:30	0:00	100	
	J	*		(Target down for entire trial)		—	
	M	2:42		5:30	1:50	100	
	A	2:42	2:42	5:30	0:54	100	100
15	T	0.45		6:00	0:45	100	
	J	1:30		6:00	3:30	67	
	M	2:00		6:00	2:15	93	
	A	2:00	2:00	6:00	3:45	67	82
16	T	0:42		6:00	3:18		
	J	5:00		4:00	2:21		
	M	6:00		6:00	4:24		
	A	6:00	*	6:00	4:45	0	0
17	T	*		2:30	0:24	84	
	J	0:48		4:00	0:48	100	
	M	*		5:00	1:36	68	
	A	0:48	0:48	6:00	2:06	78	83
18	T	*		2:00	1:48		
	J	4:00		5:00	2:36		
	M	4:30		6:00	2:12		
	A	6:00	*	6:00	4:18	0	0

*Cannot be determined

Table 2 contains the smoke screening assessment for twelve trials within the wind regime of 3.6 to 7.5 m/s. The overall effective screening assessment for this group was 99 percent.

Munition Expenditure Assessment. Table 3 shows the number of rounds KWIK predicted would obscure the entire 500 m for 6 min for two wind groups under low humidity (RH ≈ 30 percent) and under high humidity (RH ≈ 80 percent). For both humidity levels in each wind regime, the net gain or loss in munition expenditures is also indicated. In the case of trial 19, for example, the KWIK model predicted 7 munitions for low humidity and 5 munitions for high humidity, while the FM method predicted 12 munitions. This translates to a munition savings for KWIK of 42 percent and 58 percent, respectively. Munition expenditures cannot be obtained from the FM for windspeeds outside the ranges shown in table 3.

TABLE 2. SMOKE SCREENING ASSESSMENT, WINDSPEED: 3.6 - 7.5 M/S

Trial No.	Target	Initial Time To Screen Target (min: sec)	Approx Build Up (min: sec)	Target Schedule (min: sec)	Time Target Observed (min: sec)	Effective Screening (%)	Mean Effective Screening (%)
5	Tank (T)	*		4:30	0:00	100	
	Jeep (J)	*		2:30	0:12	92	
	Moving (M)	1:12		5:00	0:12	100	
	All (A)	*	1:12	5:00	0:15	95	96
6	T	2:00		3:00	2:00	100	
	J	2:00		5:00	1:18	94	
	M	1:40		3:00	1:40	100	
	A	2:00	2:00	5:00	2:00	100	98
7	T	*		2:30	0:00	100	
	J	1:00		4:00	1:15	94	
	M	*		5:00	0:00	100	
	A	1:00	1:00	6:00	1:00	100	98
8	T	*		2:00	0:00	100	
	J	1:00		5:00	0:00	100	
	M	0:30		6:00	0:30	100	
	A	0:30	1:00	6:00	0:18	100	100
9	T	1:00		3:30	0:30	100	
	J	0:30		3:30	0:00	100	
	M	1:00		5:00	0:00	100	
	A	1:00	1:00	5:00	0:00	100	100
10	T	*		3:00	0:00	100	
	J	0:24		3:00	0:24	100	
	M	0:30		5:30	0:00	100	
	A	0:24	0:30	6:00	0:24	100	100
12	T	0:18		3:00	0:18	100	
	J	1:00		5:00	0:15	95	
	M	*		3:00	0:15	92	
	A	0:18	1:00	6:00	0:18	100	97
13	T	1:00		3:00	1:00	100	
	J	1:24		6:00	1:24	100	
	M	0:48		6:00	0:57	97	
	A	1:12	1:24	6:00	1:27	96	98
19	T	0:30		3:30	0:00	100	
	J	0:30		3:30	0:00	100	
	M	1:00		5:00	0:00	100	
	A	1:00	1:00	5:30	0:18	96	99
20	T	*		3:00	0:00	100	
	J	0:42		3:00	0:42	100	
	M	0:30		5:30	0:20	100	
	A	0:50	0:50	5:30	0:50	100	100
21	T	1:30		4:30	0:00	100	
	J	1:00		3:00	0:00	100	
	M	1:25		5:00	0:25	100	
	A	1:25	1:30	5:00	0:25	100	100
22	T	0:24		3:00	0:24	100	
	J	1:00		5:00	0:00	100	
	M	*		3:00	0:00	100	
	A	0:24	1:00	6:00	0:24	100	100

*Cannot be determined

TABLE 3. MUNITION EXPENDITURE ASSESSMENT

Test No.	Field Manual	KWIK (RH 30%)	% Change (RH 30%)	KWIK (RH 80%)	% Change (RH 80%)	Wind
Windspeed: 2.0 - 3.5 m/s						
11	22	12	+45	8	+64	Qtr/Cross
14	14	8	+43	8	+43	Qtr/Cross
15	8	8	0	6	+25	Qtr/Cross
16	8	8	0	8	0	Cross/Qtr
17	8	9	-12.5	7	+12.5	Qtr/Cross
18	10	10	0	8	+20	Cross/Qtr
		Net Change	+21.4	Net Change	+35.7	
Windspeed: 3.6 - 7.5 m/s						
5	10	9	+10	7	+30	Cross/Qtr
6	10	9	+10	7	+30	Cross/Qtr
7	8	9	-12.5	7	+12.5	Crosswind
8	10	9	+10	6	+40	Crosswind
9	17	18	-6	9	+47	Qtr/Head
10	25	27	-8	15	+40	Qtr/Head
12	12	18	-50	9	+25	Quartering
13	12	15	-25	7	+42	Qtr/Cross
19	12	7	+42	5	+58	Quartering
20	11	12	-9	6	+45	Cross/Qtr
21	10	9	+10	7	+30	Cross/Qtr
22	13	12	+8	6	+54	Cross/Qtr
		Net Change	-2.7	Net Change	+39.4	

For the lower windspeed group, KWIK produced a net gain in munition savings of 21.4 percent for low humidity and 35.7 percent for high humidity. For the higher windspeed group, KWIK produced a net loss of 2.7 percent for low humidity, but a net gain of 39.4 percent for high humidity. As shown in figure 2, KWIK efficiency in munition expenditures improved as the day progressed, when compared to the FM method. This is due to steadily increasing instability from daytime heating. Under the low windspeeds (2.0 to 3.5 m/s), KWIK has the capability of finely describing atmospheric stability, while the FM method has only three gross categories. Under the higher windspeeds (3.6 to 7.5 m/s), atmospheric stability tends to remain relatively constant with daytime heating, and therefore little difference in munitions expenditures is noted in figure 3.

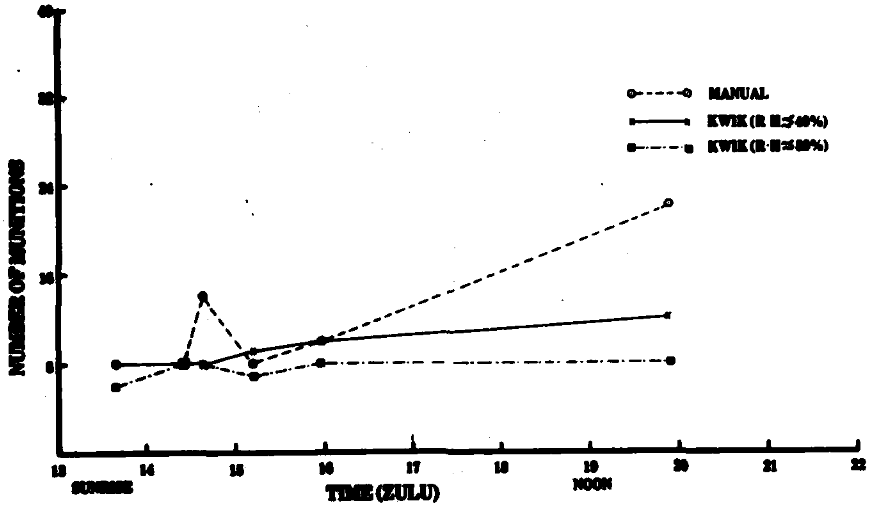


Figure 2. Munition comparison: KWIK vs field manual using Horizontal Grid met (windspeed: 2.0 to 3.5 m/s).

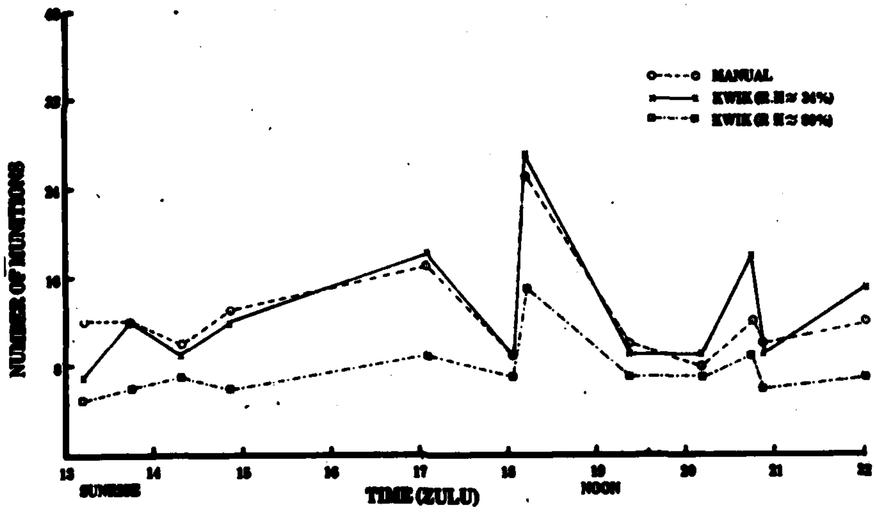


Figure 3. Munition comparison: KWIK vs field manual using Horizontal Grid met (windspeed: 3.6 to 7.5 m/s).

The alternating dash-dot curve in figures 2 and 3 represents the number of munitions KWIK calculated to be necessary to screen at a relative humidity of about 80 percent. KWIK's capability of using the hydroscopic properties of HC smoke enables more efficient use of this munition when compared to the FM method. This is demonstrated by the consistently lower munition expenditures calculated for all the plotted trials in figures 2 and 3. A net savings of 35.7 percent during the low windspeeds and 39.4 percent during the higher windspeeds could have been realized under these higher humidities. This higher humidity category (RH ≈ 80 percent) is a fairly common occurrence in Europe, especially during the predawn and early morning hours. Under these conditions, KWIK could save a significant number of smoke munitions.

Figures 4 and 5 compare munition expenditure calculations based on both the horizontal grid meteorology and the Ditto meteorology. For the low windspeeds (figure 4) KWIK consistently calculated lower munition expenditures with the horizontal grid meteorology. This calculation is not entirely unexpected, because at low windspeeds local effects caused by terrain features and solar heating tend to dominate the microscale meteorology. For the higher windspeeds (figure 5) there is no apparent mean difference between the two sets of meteorological data, although there were wide differences on any given trial between the grid meteorological data and the Ditto meteorological data. One possible explanation for this variation could be Granite Mountain, which is just a few kilometers southwest through southeast of the grid. This might have set up mountain lee waves with a southeast wind flow, resulting in much higher windspeeds at the test grid than at the Ditto Meteorological Station.

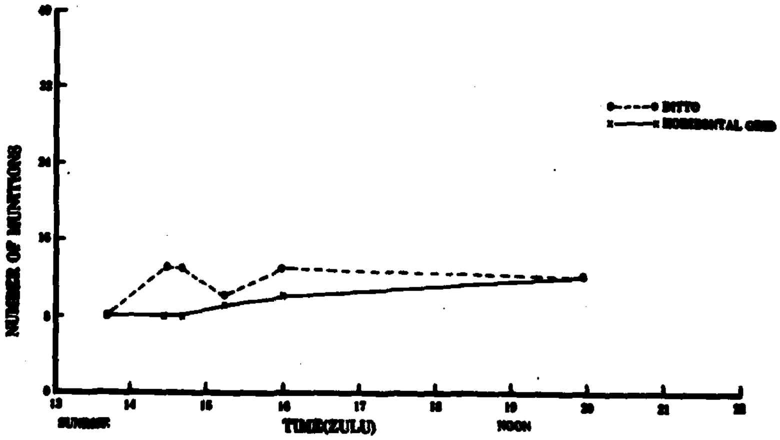


Figure 4. Meteorological comparison: Ditto vs Horizontal Grid met data (windspeed: 2.0 to 3.5 m/s).

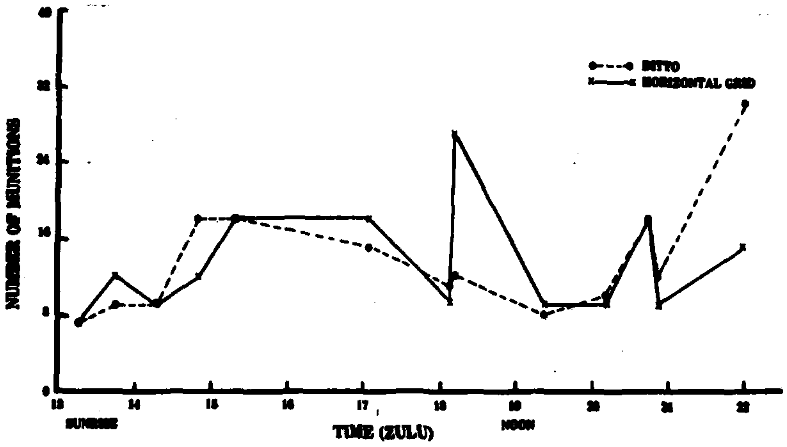


Figure 5. Meteorological comparison: Ditto vs Horizontal Grid met data (windspeed: 3.6 to 7.5 m/s).

High Wind Screening. Trials 23 through 30 were conducted at windspeeds in excess of 16 knots. Army Training Circular TC 6-20-5, entitled "Field Artillery Smoke," indicates that smoke screening at these windspeeds is unfavorable. Nevertheless, it was decided that the testing would continue as long as a successful screen could be deployed, since no other data of this type existed. As shown in table 4A, windspeeds ranged from 18 to 30 knots (8.8 to 15 m/s), with gusts to 40 knots (20 m/s). At these windspeeds a neutral stability was maintained through all trials,

TABLE 4A. WINDSPEED (WS) >8 M/S VS CALCULATED MUNITION EXPENDITURES

Trial No.	WS (m/s)	Rounds Detonated	Rounds Calculated (Grid)	Rounds Calculated (Ditto)	Rounds Calculated (RH ≈ 80%)	Direction
23	9-12	17	18	18	9	Cross/Quarter
24	8.8-12	16.75	18	18	9	Cross/Quarter
25	9.7-13	17	18	18	9	Cross/Quarter
26	8.7-12	15.75	18	18	9	Cross/Quarter
27	11.7-15.5	20	21	18	12	Quarter
28	12.8-18	25.5	27	24	15	Quarter
29	12.3-19.5	25.5	27	24	15	Quarter
30	15-20	41	45*	-	-	Quarter/Head

*Calculated during near gale

as was later verified by examining cloud behavior from photographic records. The number of munitions calculated (Grid) for the 6-min screens ranged from 18 to 45, with the higher figure calculated during a near gale with a quartering/headwind direction. Table 4A shows the number of munitions needed to successfully form a screen (from a low of 18 to a high of 27) for trials 23 through 29. The meteorological input from the Ditto Meteorological Station indicated identical munition expenditures during the morning hours, but somewhat lower amounts during the afternoon hours when the winds were higher. Since the higher winds were experienced at the horizontal grid, probably due to mountain lee waves, it was not surprising to see the slight difference in munition expenditures as calculated from data taken at the two meteorological stations.

The smoke screen characteristics for the high wind cases are indicated in table 4B. The build-up time is defined as the time, after T-0, required to obscure all targets from the observers' view. The duration of the screen is the time period from initiation of a test to the instant when a target became visible to one or more of the observers. The total effective screening time is the total time that all targets were continually screened from all observers.

TABLE 4B. SCREEN CHARACTERISTICS

Trial No.	Build Up Time (s)	Duration of Screen (min: sec)	Total Effective Screening Time (min: sec)
23	30	6:40	6:10
24	30	6:50	6:20
25	40	6:40	6:00
26	42	6:45	6:03
27	15	6:55	6:40
28	28	6:50	6:22
29	30	6:40	6:10
30	15	6:45	6:30
Mean	28.75	6:46	6:17

For the eight trials, the mean build-up time was 28.75 s, with a mean effective screening time of 6 min and 17 s. In all cases, once the screen had formed there were no apparent holes until the screen began to break up at the end of the trial. It was surprising to note that the best screens in quality and duration, as judged by observer and photographic data, occurred during the higher winds.

The similarity of the munition expenditure calculations from both the Ditto Meteorological Station and the horizontal grid met station would seem to indicate that under certain synoptic scale events, the target meteorology is similar to the meteorology several kilometers away. These types of large scale weather systems are not unusual in Europe, especially during the winter months. Another common feature during the winter storms is high relative humidity (RH ≈ 80 percent). Results using such a high relative humidity are indicated in table 4A. All other meteorological parameters are identical. A reduction in munitions of 47.39 percent over the cases with lower humidities illustrates the wide variation possible under varying ambient moisture conditions. This variation is important, considering that the FM method does not have the capability to screen under high winds nor to use the ambient moisture to reduce expenditures under high humidities.

4. CONCLUSIONS

Screening Effectiveness. In this initial phase of testing, KWIK has demonstrated that it not only is more efficient in munition utilization than the FM method, but that it also has the capability to calculate munition expenditures under meteorological conditions that the present FM method considers impractical. For the low windspeed or marginal screening category (2.0 to 3.5 m/s), two trials were unsuccessful in forming a screen because of low windspeeds and extreme variability of wind direction. Since calculations for these same two trials using the FM method produced identical munition expenditures, both techniques failed to successfully screen under these meteorological conditions. The remaining four trials in the low wind category produced a mean effective smoke screen during 88.3 percent of the screen duration time. The few instances during which a target was visible were almost always caused by significant changes in windspeed and wind direction during the course of a trial.

There were twelve trials conducted under favorable screening conditions, with windspeeds ranging from 3.6 to 7.5 m/s. These trials produced a 99-percent mean effective screen. All of these screens would have denied acquisition of a target for the duration of the screen.

Munition Expenditure Comparisons. Phase I was conducted under dry conditions, with an average relative humidity of only 30.5 percent. These dry conditions are important, because HC is a hydroscopic smoke whose screen capabilities are greatly enhanced under the higher humidities that are frequently found in Europe. For the marginal screening category, KWIK used 21.4 percent fewer munitions than the FM method. However, under high humidities (RH ≈ 80 percent), KWIK could have saved 35.7 percent of the rounds that the FM method required for the same missions.

In the favorable screening category KWIK used 2.7 percent more munitions than the FM method required. However, under high humidities (RH ≈80 percent), KWIK would have produced a 39.4 percent savings in munitions expenditures. The failure to incorporate the relative humidity effects into the FM method clearly causes an excess expenditure of smoke rounds under the higher humidity conditions.

Target Area Meteorology. Under marginal screening conditions, the winds are variable in both space and time. Even under the relatively uniform terrain of DPG, use of the target area meteorology produced a savings in munition expenditures, as shown in figure 4. As windspeeds increase, local wind circulations disperse and the general flow becomes more uniform. During the transition period between low and high windspeeds, tremendous variability can exist over a spatial distance of only 10 km (figure 5). In the wintertime European scenario, major storms covering hundreds of kilometers are quite common. Many of these storms are associated with windspeeds high enough to preclude the necessity of knowing target area meteorology to perform a mission. (This assumes that the terrain features do not dominate the target area meteorology.) Under weaker wind regimes, which occur in Europe during the summer season, local wind circulations would make the availability of target area meteorology desirable or even necessary for the completion of a mission.

High Wind Screening. One of the surprises of Phase I testing was the discovery that it is practical to screen a target during high winds. Eight trials were conducted under high wind (8.0 to 15.0 m/s) or "unfavorable" screening conditions, with a 100-percent mean effective screen. All these screens obscured all targets for more than the required 6 min, as seen in table 4B.

There are several possible reasons for the successful screening at high windspeeds. One plausible explanation is that the vegetation flattens as the windspeed increases, lowering the effective roughness length. This lowering of the roughness length would change the dispersion parameters, which are critical to the correct calculation of munition expenditures. Examining table 4A, it is noted that the actual number of munitions that successfully detonated was always less than the number calculated. Since all of the screens were successful, this overprediction of needed munitions could be due to wind-modified roughness length. Another possibility is the homogeneity of the terrain at DPG. It is possible that for heterogeneous terrain or terrain with extensive vegetation, such as a forest, the increased turbulence at higher windspeeds would preclude the use of smoke. A third possibility is a change in the efficiency of continuous burning smoke munitions during high winds. An increase in the oxygen available to the munition could conceivably improve the efficiency of such munitions.

Regardless of the reason or combination of reasons for the high wind smoke screening, further investigation is clearly warranted. If future experiments confirm that smoke screening at high windspeeds is feasible, then a change in doctrine would be indicated. This could give friendly forces an advantage in future confrontations using smoke.

6. REFERENCES

1. Pasquill, F., Atmospheric Diffusion, 2nd edition Halsted Press, division of John Wiley & Sons, Inc., New York, 429 pp, 1974.

2. Turner, D. B., "A Diffusion Model for an Urban Area," J Appl Meteorol, 3:83, 1964.

3. Smith, F. B., "A Scheme for Estimating the Vertical Dispersion of a Plume from a Source Near Ground-Level," unpublished Meteorological Office note, 1973.

4. Downs, A. R., A Review of Atmospheric Transmission Information in the Optical and Microwave Spectral Regions, Report 2710, US Army Ballistic Research Laboratory, Aberdeen Proving Ground, MD, 1976.

5. Gordon, Jacqueline I., Daytime Visibility, A Conceptual Review, AFGL-TR-79-0257, Visibility Laboratory, Scripps Institute of Oceanography, La Jolla, CA, 1979.

6. Schuppert, G. T., "Visibility Concepts and Measurement Techniques for Aviation Purposes," DOT-TSC-FAA-71-25, Transportation Systems Center, Cambridge, MA, 1971.

7. Gifford, F. A., "An Outline of Theories of Diffusion in the Lower Layers of the Atmosphere," Meteorology and Atomic Energy, D. Slade, editor, US Atomic Energy Commission, Washington, DC, 1968.

8. Umstead, R. K., R. Pena, and F. V. Hansen, KWIK: An Algorithm for Calculating Munition Expenditures for Smoke Screening/Obscuration in Tactical Situations, ASL-TR-0030, US Army Atmospheric Sciences Laboratory, White Sands Missile Range, NM, 1979.

9. JTCG/ME, "Summary Tables of Estimated Ammunition Expenditures to Establish and Maintain Smoke Screens," unpublished manuscript.

10. Field Artillery Smoke, Training Circular 6-20-5, US Army Field Artillery School, Fort Sill, OK, 1975.

11. Modern Battlefield Cannon Gunnery, Field Manual 6405, Headquarters, Department of the Army, Washington, DC.

HARDENING OF ARMORED VEHICLE SUSPENSION COMPONENTS

HUBERT COMMINGE, MR.
US ARMY MOBILITY EQUIPMENT R&D COMMAND
FORT BELVOIR, VA 22060

INTRODUCTION

Mechanically-emplaced and air-delivered land mines have been introduced into the modern battlefield thus compounding the problem of maintaining armored-vehicle mobility. Land mines and mined areas will be encountered by U.S. Forces throughout the battlefield during offensive, defensive, and rear-area support operations. In tracked vehicles, the vulnerable components are those associated with the suspension system and, most particularly, the tracks and roadwheels. This is a critical problem area; a small 3-pound explosive charge can cause a mobility failure on all tracked vehicles. The major damage resulting from such an explosion is the loss of a track section, resulting in an immobile vehicle. Since minefields are usually covered by artillery and anti-tank weapons, an immobilized vehicle can be destroyed quickly.

Historically, warfare scenarios and deployment techniques have limited the use of land mines to certain well-defined situations to impede or channel opposing forces. Advance knowledge of minefield locations could permit effective countermeasures using mine-clearing equipment such as mine-clearing rollers mounted on lead vehicles, projected explosive line charges, or, depending on mine density, bulling through with some sacrifice. The advent of offensive mining through air- and artillery-delivery capabilities has introduced a new tactical situation that requires mobile units to proceed across random-laid minefields or mine areas.

Vulnerability analysis has predicted severe losses of armored combat units in all mine environments. This prompted investigations for providing countermine hardened suspensions for armored vehicles. The following guidelines were developed for hardening tank track shoes and roadwheels thus providing enhanced survivability:

a. Suspension components should be hardened to demonstrate enhanced blast survival from the explosion of a undirectional shallow-buried charge.

b. Hardened components should not exceed the weight of current suspensions.

c. The M60 tank tracks and roadwheels should be the baseline suspension components.

d. The vehicle should be mobile after a single encounter and should be able to "limp-off" the minefield and preclude a complete kill (K-Kill).

SUSPENSION FAILURE ANALYSIS

Current tank track shoes and roadwheels are severely damaged when subjected to an explosive charge, even a small (3-pound) track-cutting explosive charge. An analysis was conducted to determine the mode of failure in order to develop the protection mechanism.

In explosive charge-target geometries, where the explosive is in direct mechanical contact with the target, the predominant failure mechanism is shear caused by a large-velocity gradient. This can be illustrated qualitatively by Figure 1.

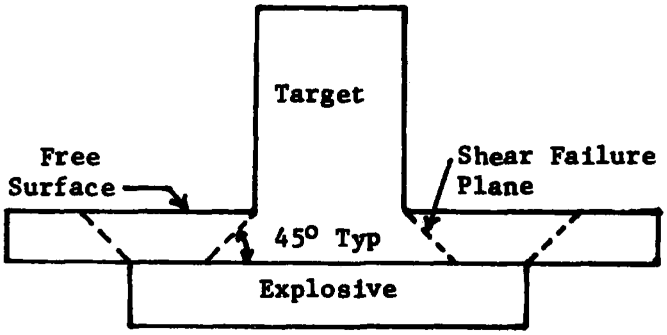


Figure 1. Arbitrary charge/target geometry.

In this arbitrary target, the explosive, simulating an air-scatterable mine, is in direct contact with the target. Detonation of this charge causes a high-pressure shock wave to be propagated into the target material adjacent to the charge. The interface between explosive and target moves upward at a high velocity. Material along the bottom edge of the target, not in direct contact with the explosive charge, receives no velocity wave and, therefore, remains initially at rest. This large velocity differential between material in contact with and material not in contact with the charge, induces a strong shear stress in the target material. Since the compressive stress is essentially one-dimensional, a shear failure is expected to propagate from the corners of the charge upward along the dashed lines. A similar mechanism also causes shear failure along the dashed lines emanating from the internal corner of the target's top surface. The initial pressure-velocity wave propagates upward until it reaches the free surface; at this time a rare-faction shock reflects from the free surface and imparts a velocity which is approximately twice that of the incident velocity. A velocity gradient exists between the material under the free surface and that where no free surface is encountered; therefore, a shear failure is expected.

Verification that a shear failure can occur around a discontinuity in a target structure is illustrated in Figure 2 below.

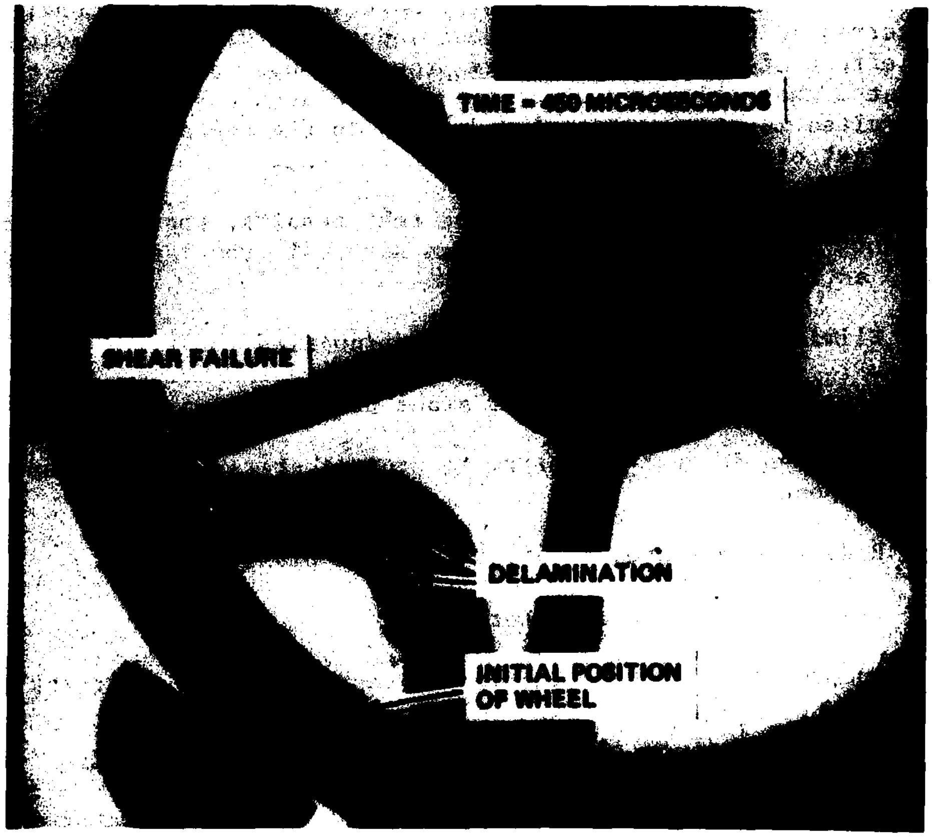


Figure 2. Dynamic X-ray of wheel deformation.

This doubly exposed radiograph shows a wheel section before and 450 microseconds after initiation of an explosive charge. The wheel was specially constructed for this test and incorporated chosen discontinuities. The wheel rim consisted of weldments of laminated bands and alternating layers of steel and aluminum, with the outermost band being aluminum. The explosive charge, as shown in the radiograph exposure, was placed directly against the wheel to simulate an air-scatterable mine. A shear failure starting from the spoke-rim joint is clearly demonstrated in the 450-microsecond X-ray exposure.

Tests similar to that described above were conducted with different tank roadwheel and track-shoe configurations to locate critical failure points in component structures and to determine how these failures could be controlled or eliminated. Discontinuities within the mine-target environment were the key to component failure. First, almost all observed failures occurred around discontinuities resulting in large-velocity gradients between adjacent portions of material; second, discontinuities that occurred in the structure without inducing a velocity gradient often had the effect of inhibiting failure propagation. In addition, tests demonstrated that reduction of target surface area normal to the explosive blast resulted in a proportional reduction in the overall load delivered to the target structure.

From the roadwheel and track-shoe test results, the protection requirements for tank suspension system survival from landmines were developed as follows:

a. Eliminate sharp geometric discontinuities.

b. Use laminations to inhibit crack propagation.

c. Pre-plan fracture geometry to protect critical tank suspension components.

d. Reduce normal area adjacent to the explosive charge.

e. Use energy-absorbing materials.

f. Incorporate blast deflection in tank roadwheel and track design.

ROADWHEEL AND TRACK DESIGN

The major disadvantages to the current tank roadwheel and track shoe structure, when subjected to a track-cutting explosive charge, are the large flat surface areas presented to the explosive blast source;

in addition, the components contain sharp geometric discontinuities. The large area absorbs the full brunt of the explosive blast and is not conducive to dissipating or deflecting the shock wave. Further, the metal components transmit the large-velocity gradient to other suspension components such as roadarms and torsion bars and cause additional damage.

The new suspension configuration encompasses a single tank roadwheel in lieu of the present dual-roadwheel system and a track shoe having a semi-circular groove for maintaining proper track alignment and interface with the roadwheel. This single roadwheel allows a large amount of the explosive blast to flow by and dissipate quickly; hence decreasing/eliminating the damage to the roadwheel and other suspension components such as the roadarm, shock absorbers, and torsion bars.

The hardened roadwheel internal structure consists of seven fiberglass radial support rings with an outer ring encapsulated in an elastomeric material as shown in Figure 3.

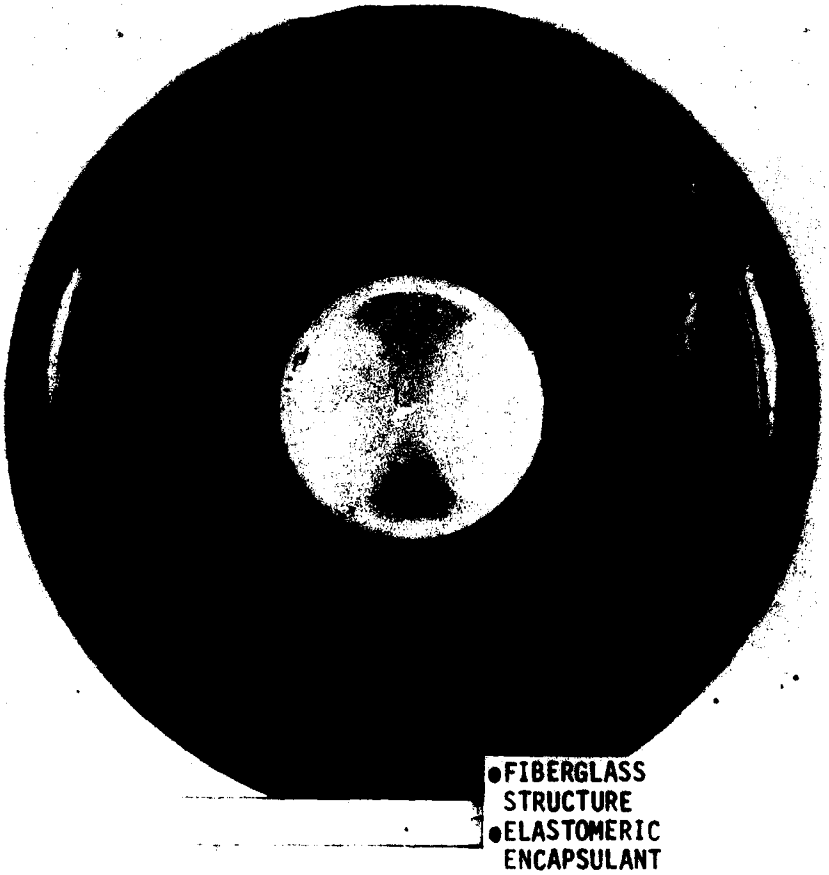


Figure 3. Mine-resistant tank roadwheel.

The support rings are fabricated of E-glass tape bonded with EPON 828 epoxy and Z-modified polyamine; the cure procedure consists of hardening at room temperature and 4 hours at 275°F. The encapsulant is Adiprene APR-882 polyurethane and MOCA-cured.

The mode of deformation for the roadwheel structure under load from an explosive blast occurs initially at the single point of tangency between the outer rim and one of the inner rings. As the explosive blast load is applied, both rim and inner ring deform so that the area of mutual contact increases continuously. Since rim-ring contact area increases continuously during loading, structural discontinuity does not occur at a single point and the large-velocity gradient is spread over a wider area minimizing and/or eliminating structural damage depending upon explosive charge size.

The hardened track shoe was developed to provide the required structual integrity, proper floatation, and minimum area of primary structure adjacent to the explosive blast. The shoe consists of two metallic links with accommodating pins and bushings, a metallic yoke, and a non-metallic track-shoe body as shown in Figure 4.

Figure 4. Mine-resistant tank track shoe.

The links, yoke, and pins manufactured of 4340 steel and hardened to a Rockwell C43 provide necessary structural integrity for both mobility and explosive blast and are configured to minimize the area of primary structure. The primary purpose of the yoke, besides adding to the structural integrity of the track shoe, is to eliminate the latitudinal separation of the track links under explosive blast. Earlier testing of hardened track shoes without yokes resulted in the breaking of large numbers of non-metallic track-shoe bodies because of excessive longitudinal and latitudinal motions of the track when subjected to an explosive blast. The track-shoe body consisting of Adiprene L-315 (Shore D-70) molded to the yoke, provides the necessary vehicle floatation and traction and is sacrificed during the explosive blast; this loss allows for quicker dissipation and/or venting of the shock wave. The track-shoe design is such that the links are inserted sideways into the opening of the yoke and rotated 90 degrees to lock in place; this design eliminates the currently used track-shoe end connectors and accessories. Further, the center guides are completely eliminated; track guidance is achieved by the semi-circular groove of the track shoe body that allows better wheel traversing for terrain contour.

ENERGY ABSORBING MATERIALS

Detonation of a landmine releases a large quantity of energy in a very short period of time, creating a large moment of inertia on the track and roadwheel subsystem. Prior to destruction, the roadwheel imparts a moment to the roadarm assembly which is transmitted to the torsion bar, causing catastrophic failure of these components. In addition, a large antitank mine containing 15 pounds or more of explosive will remove the torsion bar housing from the tank hull. Materials, other than metals, capable of absorbing some of the explosive energy were required to decrease the overall transmitted inertia and decrease suspension component destruction.

Different energy-absorbing materials were analyzed; the most promising material compositions were subjected to explosive blast tests. It was obvious that full-scale explosive testing of all promising energy-absorbing materials and design configurations would be prohibitive; therefore, a low-cost testing procedure was developed to eliminate undesirable composite materials. The testing was accomplished by fastening the energy-absorbing material to a rigid I-beam with one end pinned to a concrete foundation (as shown in Figure 5). The beam was free to rotate about its pin joint, and rotation was measured by means of a mechanical scribe. Beam mass and moment of inertia were determined prior to each test so that energy collected by the composite material sample and transmitted to the beam could be calculated from measured beam rotation. A baseline material was required to compare the energy dissipated or absorbed by each composite material sample; steel was chosen as the baseline material, with the assumption that it would not absorb and transmit the energy directly to the beam.

TEST BEAM
SPECIMEN
5 POUND CHARGE

Figure 5. Energy absorption test set-up.

The table gives a partial listing of composite materials tested for the tank track shoe. The vehicle must be capable of traveling after a mine detonation; therefore, the composite material should dissipate 100 percent of the energy to eliminate damage to other suspension components and yet remain intact to allow for vehicle floatation. For this reason, the composite materials having the highest percentage value of energy dissipated with the least damage were employed in full-scale testing on a combat tank.

Results of Energy Tests of Track Composite Materials

	MATERIAL	ENERGY DISSIPATED (%)	DESCRIPTION OF BLAST DAMAGE
THERMOSETTING	Epoxy-Epon 828/Catalyst Y	60	Partially fragmented
	Urethane-Uralite 3121S	18	Minimal
	Epoxy-Epon 828/Catalyst Y, Silica-Reinforced	53	Partially fragmented
	Epoxy-Epon 828/Catalyst Y, Steel-Reinforced	58	Fragmented
	Non Rigid Epoxy-Epon 282/DDSA	61	Fragmented
	Polyster-Glass-Reinforced	54	Pulverized
	Epoxy-25% Keular-¼ inch	64	Partially fragmented
THERMO-PLASTIC	High-Density Polyethylene	64	Eroded
	Polycarbonate--Unfilled	59	Minimal
	Polycarbonate--40% Glass Fibers	63	Fragmented
	Polyurethane--40% Glass Fibers	59	Partially fragmented
	Steel	0	Totally Destroyed

EXPLOSIVE TESTS

Both the present dual-roadwheel tank suspension and single-roadwheel hardened suspension systems were subjected to numerous full-scale explosive tests. Tests were conducted with operational and simulated landmines containing 5, 13, and 22 pounds of explosive. Most foreign landmines contain cast TNT; therefore, all simulated explosive charges were fabricated with TNT.

The 5 pound explosive charge, simulating an air-scatterable mine, was placed in direct contact with the tank track as would occur in an operational situation. The present dual-roadwheel suspension system sustained the following major damage, causing a mobility kill:

a. Dual roadwheel completely demolished, whether steel or aluminum.

b. Two track shoes completely destroyed.

c. Broken torsion bar at the first station.

d. Broken adjusting link at roadarm connection.

The damage sustained by the hardened suspension from the same explosive charge was negligible, clearly demonstrating the energy absorption of the chosen materials. As a matter of fact, the tank continued its mobility mission without degradation. The damage was:

a. Loss of two or three track shoe plastic bodies (designed to be sacrificed during explosion).

b. Very slight cracking/peeling of roadwheel encapsulant.

The next phase of testing was conducted to determine damage from off-the-shelf antitank mines. The explosive charge contained 12.6 pounds of cast TNT and was encased to simulate typical Warsaw Pact landmines. The simulated mines were covered with 2 inches of soil. The dual-roadwheel suspension consisted of steel roadwheels and T-142 track. The damage from the explosive blast is shown in Figure 6.

Figure 6. Damage to present suspension system from 12.6 pounds of TNT.

The damage, as shown above, was extensive and resulted in a mobility kill as follows:

a. Complete destruction of three track shoes.

b. Complete removal of the first dual roadwheel from the spindle (the wheels were sufficiently damaged to prohibit reuse).

c. Broken torsion bar at the first station.

d. Broken adjusting link.

e. Loosening of first and second stations torsion-bar housing caused by hull mounting-bolt shear.

f. Removal of roadarm from housing and rotating forward of vehicle.

The following damage to the hardened suspension (Figure 7) was much less severe and did not result in a mobility kill:

a. Cracking of the roadwheel encapsulant.

b. Loss of three track-shoe plastic bodies (designed to be sacrificed during explosion).

c. Loosening of adjusting link.

d. Loosening of first-station torsion-bar housing.

Loosening of the idler-wheel adjusting link eliminated the track tension, but since the track links remained undamaged, the vehicle was able to proceed under its own power with mobility degradation.

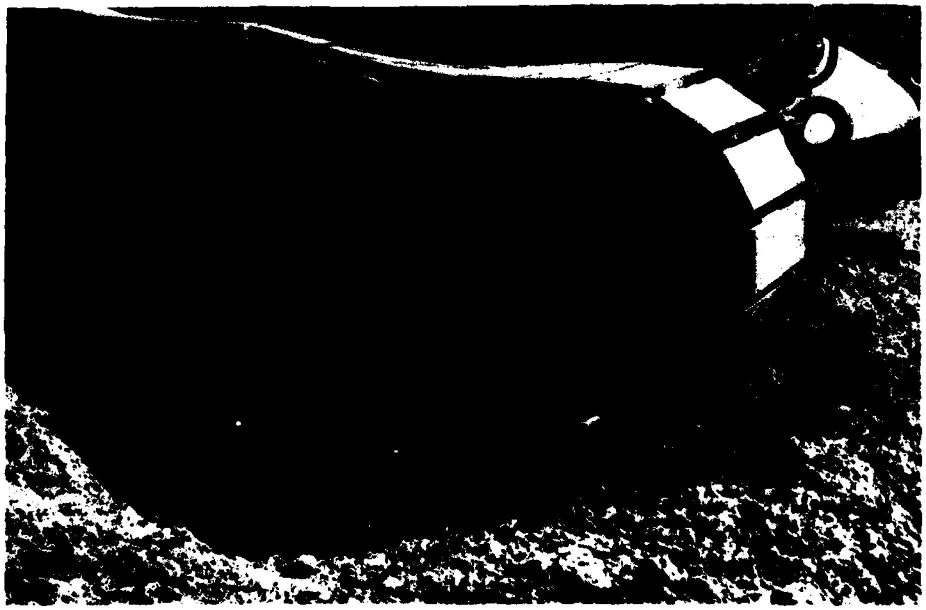


Figure 7. Damage to hardened suspension system from 12.6 pounds of TNT.

The final phase of testing was conducted using 22 pounds of explosive; this equates to the largest off-the-shelf antitank mines. The present dual-roadwheel suspension system was not tested because the damages would have been greater than previously encountered with the smaller explosive charges. The hardened suspension suffered a large amount of damage, but not sufficient to cause a mobility kill. The damage was as follows:

a. Roadwheel encapsulant is seriously damaged and inner ring structure is slightly cracked.

b. Loss of 6 to 8 track-shoe plastic bodies.

c. Loosening of adjusting link.

d. Breakage of 3 to 4 track-shoe steel yokes.

e. Loosening of both the roadarm and torsion-bar housing caused by shear of mounting bolts.

Mobility was greatly reduced, but the vehicle still proceeded under its own power. The track metal links and pins were not damaged and could engage the tank drive sprocket.

CONCLUSION

The goal of the hardening program was to demonstrate that damages to tank suspension systems from the detonation of a landmine could be minimized and, thus, decrease mobility kills. Through redesign and judicious material selection, a new suspension concept was developed and demonstrated successfully, thus eliminating armored vehicle mobility kills and allowing the vehicle to continue its intended mission or to be repaired for the next battle.

RESPONSE OF MX HORIZONTAL SHELTER MODELS TO STATIC AND DYNAMIC LOADING (U)

VAN T. COST, MR.* AND CAYLE E. ALBRITTON, MR.
U. S. Army Engineer Waterways Experiment Station
Vicksburg, Mississippi 39180

INTRODUCTION

A concept considered for basing the Missile X System was to emplace the missile in a buried, horizontal shelter. In support of the MX Horizontal Shelter Component Test Program, the U. S. Army Engineer Waterways Experiment Station (WES) developed and performed a program of static and dynamic model tests so that response data on specific structural details and shelter components could be obtained. The objectives of the test program were to evaluate the load capacity merits of various shelter soil-structure design concepts and to optimize certain parameters, while providing a design team with timely data necessary for structure survivability verification and construction economy comparisons. Static load vs. damage data from such tests provide necessary input for dynamic analyses, while dynamic test data help verify those analyses and the applicability of static data in predicting dynamic response.

The test articles were all 1/7.25-scale models of a cylindrical concrete shelter segment, which were assembled with thrust-restraining hardware and buried horizontally in a soil material simulating that found in a typical siting area. The first phase of the program consisted of static tests on 17 cylinder models having a number of design parameters including backfill stiffness, depth of burial, reinforcement type, liner thickness, shear stud spacing, and wall thickness, as well as the modeling of such construction details as breakout joint sections and inspection panels. In the course of these initial tests, designs of the shelter system were developing and being modified in response to test data and other factors. Upon completion of the initial 17 tests comprising Phase I, a specific structure design was adopted. A likely combination of cylinder wall thickness, backfill stiffness, and removable panel detail was chosen as the baseline for a second phase of eight static tests, in which the responses of this baseline design and certain variations thereof were examined. The baseline model structure was also tested dynamically to verify the similarity of static and dynamic response modes. The response of each model was monitored with electronic instrumentation.

This paper discusses the Phase II static tests and the associated dynamic verification tests, and summarizes their results. Conclusions concerning the relative merits of the selected wall thickness, burial depth, backfill stiffness, and opening detail for the baseline shelter design are presented.

DESCRIPTION OF TEST ARTICLES

All structures discussed herein were cylindrical concrete shells with sheet steel liners and welded shear studs. The cylinders had inside diameters of 24 inches and lengths of 48 inches. Outside diameters varied as prescribed by the wall thicknesses of 1.8, 2.4, 3.0, and 3.6 inches. Reinforcement, other than steel liners and shear studs, consisted only of a grid of wire in the outer face which modeled the prototype's temperature reinforcement. The baseline cylinder had a 3.0-inch wall thickness.

Most of the cylinder models included a removable panel in the crown section which was a model of the required inspection panel opening. This panel consisted of matching Z-shaped metal frames with imbedment dowels welded along the edges. Other construction details of the panel matched those of the adjacent structure section.

STATIC TEST PROCEDURES

All shelter model sections and appurtenant hardware were constructed at WES using materials conforming to invisioned prototype specifications. Reuseable concrete forms were machined from rolled steel for close tolerances and a smooth concrete finish. A single inner form designed for disassembly from within a cast cylinder was used for all models, since inside diameters remained constant. One outer form and one aligning baseplate were built for each wall thickness of cylinder model to be tested, so that each model of a given wall thickness could be cast in the same form assembly.

Construction of a cylinder model was begun with the fabrication of the inspection panel frame. The frame was assembled with a steel liner, shear studs, dowels, reinforcement, and appropriate instrumentation gages and gage mounts, and was mounted on a machined mandrel for casting (Figure 1). The mandrel was attached to a vibration table during casting to assure concrete consolidation. Transparent acrylic rolled to the proper curvature was fastened to the panel to serve as outer formwork. After 24 hours, the completed panel was removed from the mandrel and fastened to the inner cylinder form, along with the steel liner and welded studs, outer reinforcement, and instrumentation gage mounts (Figure 2). Outer formwork was then assembled (Figure 3), and the cylinder cast in a vertical

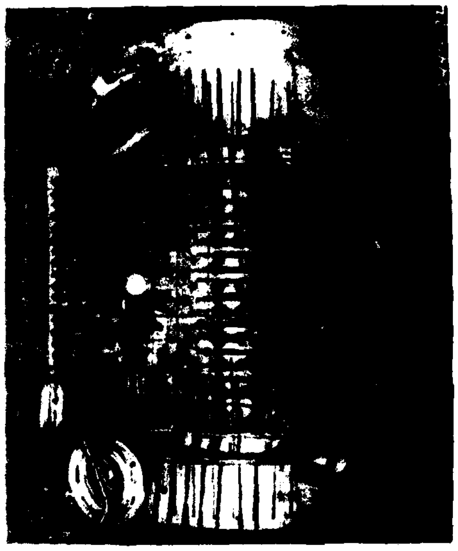


Figure 1. Inspection panel ready for casting.

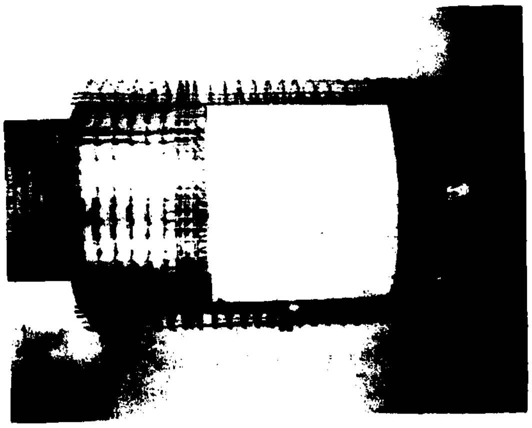


Figure 2. Model reinforcement and panel on inner form.

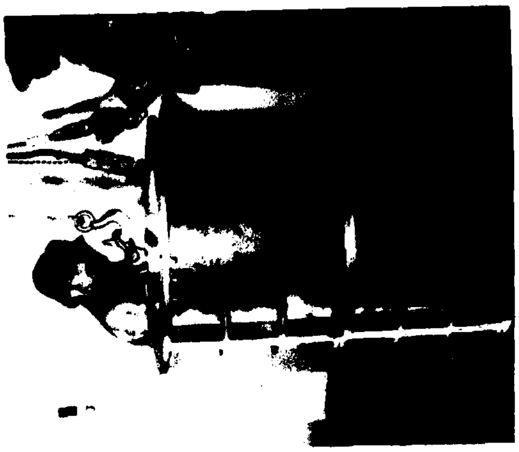


Figure 3. Outer formwork.

orientation. Concrete consolidation was accomplished using pencil vibrators carefully threaded into stud spacings marked along top flanges of the formwork.

After removal of formwork, a model cylinder was cleaned, instrumented, and assembled with the thrust-restraining hardware (Figure 4). This hardware consisted of 3-inch thick steel end plates and an internal thrust column designed to carry all of the axial loads exerted by soil pressure normal to the end plates. Rings of 1/2-inch-thick rubber were sandwiched between the model structure and each end plate to prevent the intrusion of soil without transferring significant axial load. These hardware were not considered test articles.

A specially designed loading chamber approximately 6 feet in diameter was used for the static tests. The steel test chamber was lined with 2 layers of greased polyethylene to reduce the effects of sidewall friction. Virgin soil was placed and compacted, for each test, to the elevation of the cylinder invert, where a leveled strip of sifted soil was prepared for bedding the structure. After careful positioning, the backfill process was continued by hand in the areas near the cylinder invert, to assure uniformity. Backfill density and moisture content were systematically monitored during backfilling using both sand cone and nuclear measuring device techniques (Figure 5). Two regions of soil were delineated in the test chamber: an in-situ zone having stiffness characteristics of the undisturbed soil in a prototype siting area, and a backfill zone of lesser density representing excavated and recompacted soil. Figure 6 illustrates the geometry of an assembled cylinder model and soil in the test chamber.

Water pressure from a rate-controlled pump served as the loading medium in the static tests. Water pressure was sealed from the backfill surface by a rubber membrane in a floating seal ring. The Central Firing Station (CFS) of the Blast Load Generator facility at WES was used as the reaction structure for the loading head-test chamber assembly (Figure 7).

Active electronic instrumentation in the static tests included strain gages on liners and outer reinforcement and structure interface pressure gages at several locations around the cylinder, vertical and horizontal diameter change gages, free-field soil stress gages, and loading pressure gages. Measurements were recorded on analog tape and later digitized for plotting. Remote-controlled in-structure cameras were also used to document response in some tests.

DYNAMIC TEST PROCEDURES

Models constructed for dynamic tests were of the baseline design and were identical to those tested statically. The four model cylinders were

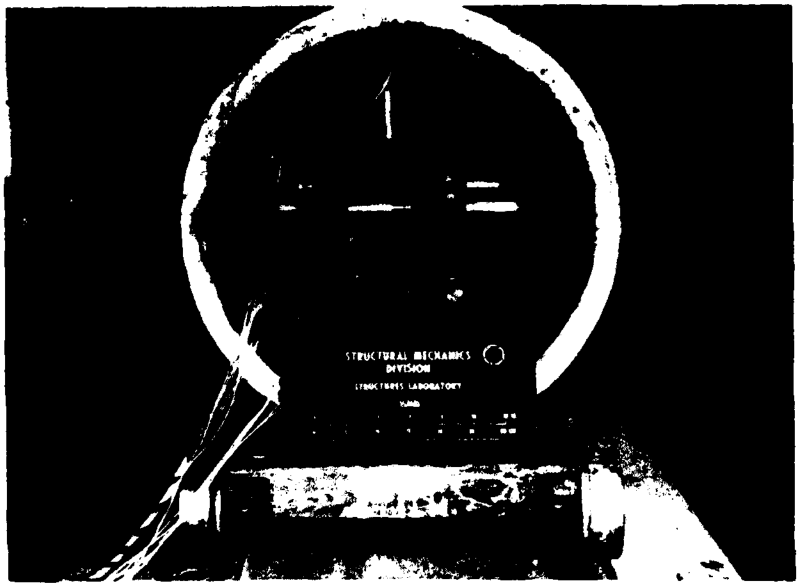

STRUCTURAL MECHANICS
DIVISION
STRUCTURES LABORATORY

Figure 4. Assembly of thrust column and instrumentation from end of model.

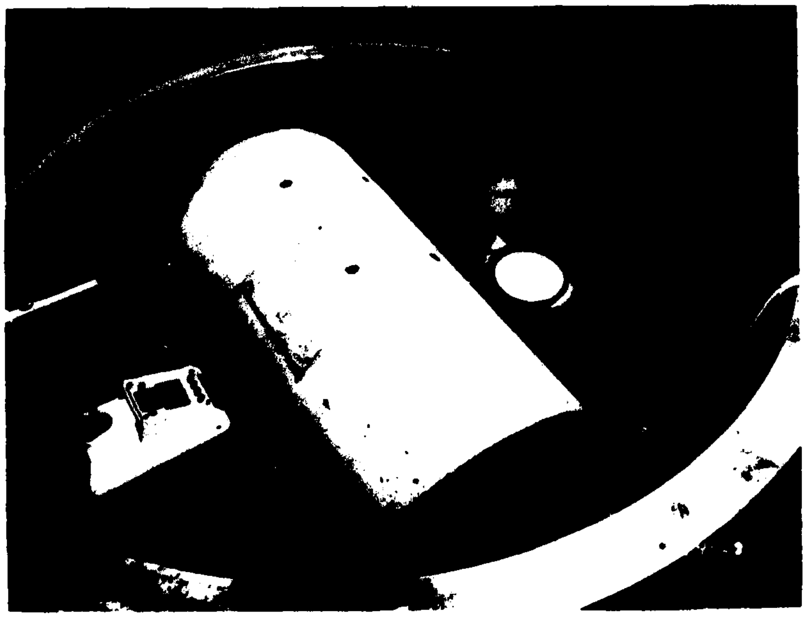


Figure 5. Backfill density control around cylinder.

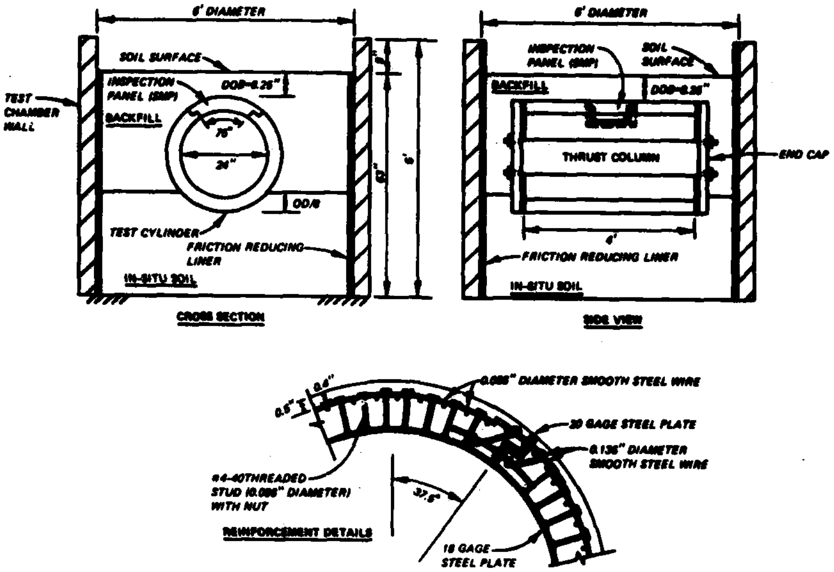

6' DIAMETER
SOIL SURFACE
TEST CHAMBER WALL
INSPECTION PANEL (SMP)
DOB=0.25"
BACKFILL
76°
24"
OD/8
TEST CYLINDER
FRICTION REDUCING LINER
IN-SITU SOIL
CROSS SECTION
6' DIAMETER
INSPECTION PANEL (SMP)
SOIL SURFACE
BACKFILL
DOB=0.25"
THRUST COLUMN
END CAP
4'
FRICTION REDUCING LINER
IN-SITU SOIL
SIDE VIEW
0.4"
0.5"
0.086" DIAMETER SMOOTH STEEL WIRE
20 GAGE STEEL PLATE
0.136" DIAMETER SMOOTH STEEL WIRE
#4-40THREADED STUD (0.086" DIAMETER) WITH NUT
37.5°
REINFORCEMENT DETAILS
18 GAGE STEEL PLATE

Figure 6. Horizontal cylinder test concept-Phase II tests.

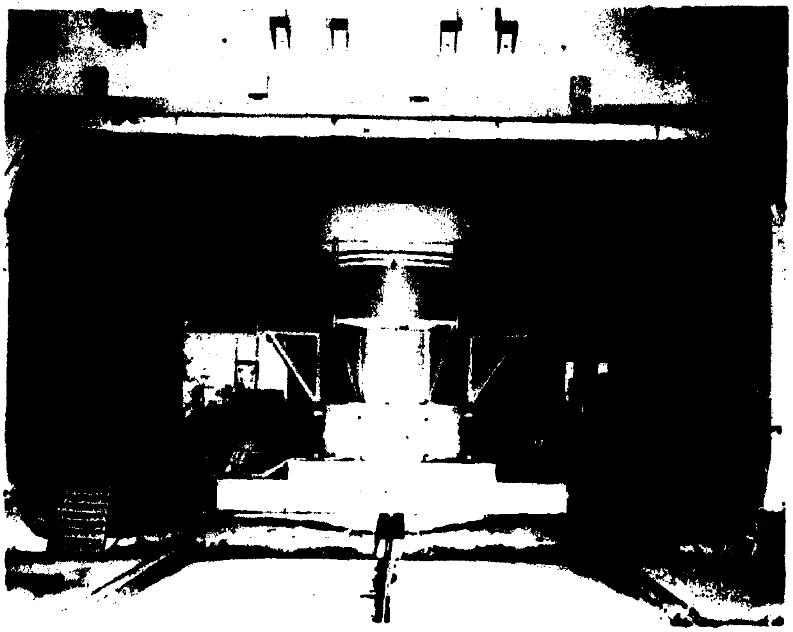


Figure 7. Test chamber and loading head inside CFS reaction structure.

instrumented and assembled in the laboratory before being shipped to the test site at the Fort Polk Military Reservation in Louisiana. Several truckloads of the simulated siting area soil used in the static tests were transported from WES to Fort Polk for construction of the test beds. The FOAM High-Explosive Simulation Technique (FOAM HEST) was used to simulate nuclear weapon overpressures.

For each dynamic test, native soil was excavated in an area 16 feet square to a depth of 8 feet, and was replaced with the simulated siting area soil. As in the static tests, soil placement densities were controlled as specified for both in-situ and backfill zones (Figure 8). The FOAM HEST explosive charge was assembled over the entire backfill surface (Figure 9), and a sand overburden of appropriate thickness was placed over the charge (Figure 10). Electronic instrumentation measurements included all those made in static tests, as well as structure accelerations and airblast pressures at the backfill surface. Figure 11 illustrates the test bed layout for the dynamic tests.

BASELINE MODEL DESIGN AND INVESTIGATED VARIATIONS

Ongoing shelter design during the first phase of the WES tests produced the baseline for Phase II. The 1/7.25-scale model of the baseline design had a wall thickness of 3.0 inches and included an inspection panel spanning 75 degrees of the crown section over a length of 12 inches as shown in Figure 6. The inspection panel frame used 20 gage steel, and the liner material was 18 gage. Shear studs and outer reinforcing wire were 0.086 inch in diameter, spaced at 1.65 inches each direction. Soil densities used for baseline design tests were 128 and 121 pounds per cubic foot (pcf), dry density, for the in-situ and backfill zones, respectively. Depth of burial was 8.25 inches, measured over the crown.

The baseline configurations of the inspection panel, liner, studs, and reinforcement represented final designs and were not varied in the Phase II tests. However, one model with no inspection panel was tested to document the panel's effects on response of the baseline design. Wall thickness effects were investigated in tests of three models other than the baseline with thicknesses ranging from 1.8 to 3.6 inches. One test incorporated a backfill density equivalent to the in-situ density to examine backfill effects. Although the 8.25-inch depth of burial was not changed in the Phase II tests, this was an investigated parameter since all of the Phase I tests were performed at depths of 3 or 6 inches, and many of those structures were otherwise similar. All dynamic test configurations were of the baseline design, and varied only in the environment overpressure and duration.

It is recognized that parametric investigations of only a single test are not sufficient to accurately quantify the effects of such variables.

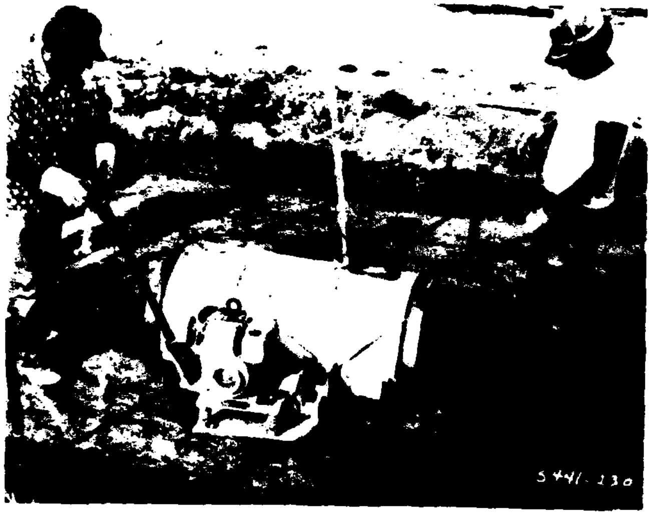


Figure 8. Dynamic test soil placement.

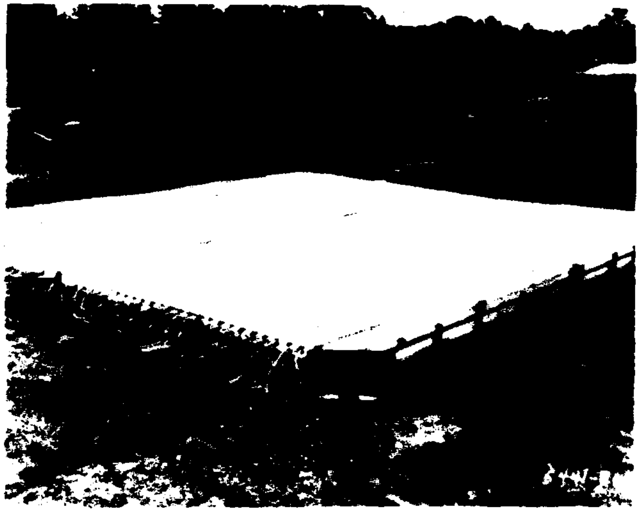


Figure 9. FOAM HEST charge.

Figure 10. Sand overburden.

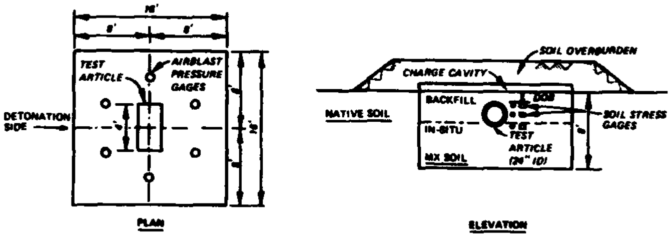

16'
8'
8'
TEST ARTICLE
AIRBLAST PRESSURE GAGES
DETONATION SIDE
16'
PLAN
SOIL OVERBURDEN
CHARGE CAVITY
BACKFILL
DOB
NATIVE SOIL
SOIL STRESS GAGES
IN-SITU
TEST ARTICLE (24" ID)
MX SOIL
ELEVATION

Figure 11. Horizontal cylinder dynamic test bed layout.

However, the primary objectives of the Phase II tests were to quantify static response of the baseline shelter design only, while identifying the trends of influence certain variations might have. Three of the eight Phase II static tests were of the baseline concept.

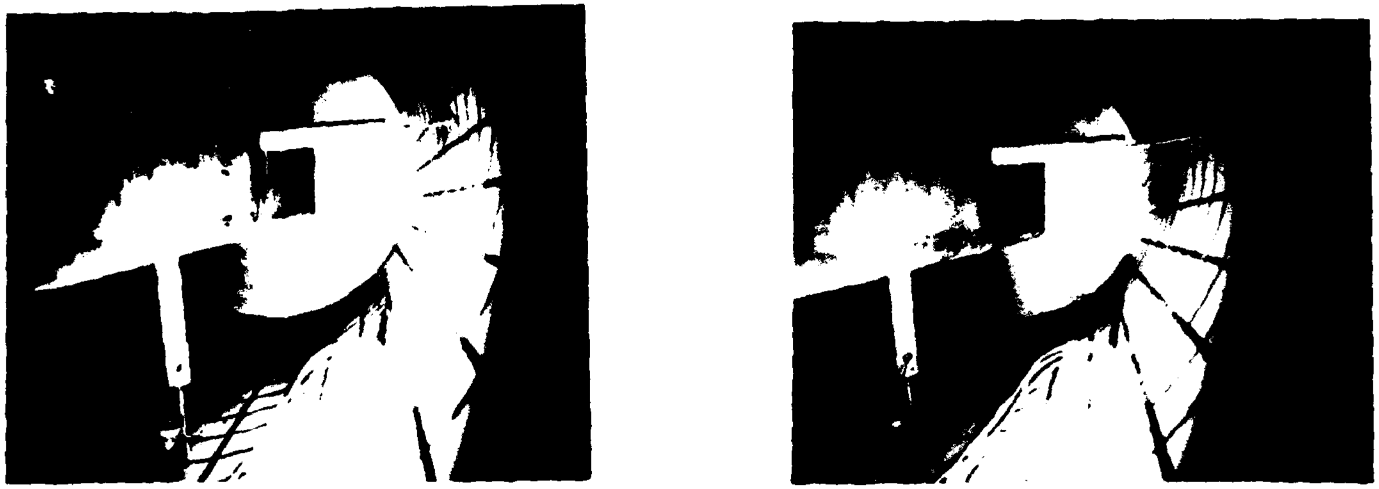
RESULTS AND DISCUSSION

As with most structural testing programs, it was necessary to specify a criteria defining a particular level of damage or, at least, some pressure level at which a test would be terminated. Due to the functional requirements of the shelter and the extreme deformation capabilities of the load-carrying soil-structure system, it was decided that static tests should be terminated prior to catastrophic failure. Since the operation of prototype missile transport systems and other hardware depended on certain interior shelter features and clear spaces, "failure" of a model structure was defined as that overpressure at which disturbances of the steel liner due to severe crushing of adjacent concrete occurred, i.e. the occurrence or initiation of buckling of the liner. Instrumentation measurements and, in some cases, midtest photography were used to identify the point during a test at which these criteria were met.

Response of the Baseline Design

The condition for the failure criteria was consistently met in the static baseline tests when severe concrete crushing occurred in areas just below cylinder springlines and buckles formed in the liners. Instrumentation and photography indicated that these buckles formed suddenly (Figure 12). The vertical diameter decrease due to ovaling deformation at the time of liner buckling ranged from 1.0 to 1.6 inches (4.2% to 6.7%) in the three baseline tests. The corresponding static overpressures ranged from about 350 to 600 psi. Comparisons of the three tests show that plots of vertical diameter change vs. surface overpressure are quite consistent, and almost linear, with slopes of about 350 to 400 psi per inch of diameter change (Figure 13). Figure 14 shows a plot of horizontal diameter elongation vs. surface overpressure. Since springline buckles formed in the areas of the deflection gage mounts in this test, their formation appeared as a disturbance on the plot, with an altered slope thereafter.

Although the response "stiffnesses" of the three baseline tests are consistent, the points at which buckles occur during loading differ somewhat with respect to diameter change and more severely with respect to surface overpressure. It is suspected that slight differences in concrete compressive strengths and local variations of backfill densities may have contributed to the range of response magnitudes in which defined failure mechanisms initiated. Measured densities in the backfill areas of the three baseline tests ranged from 119.3 to 122.6 pcf, and concrete compressive strengths at test date ranged from 6580 to 7090 psi.



a. Prior to loading.

b. 350 psi surface pressure.

c. 400 psi surface pressure.

d. 500 psi surface.

Figure 12. In-test photographs showing damage progress.

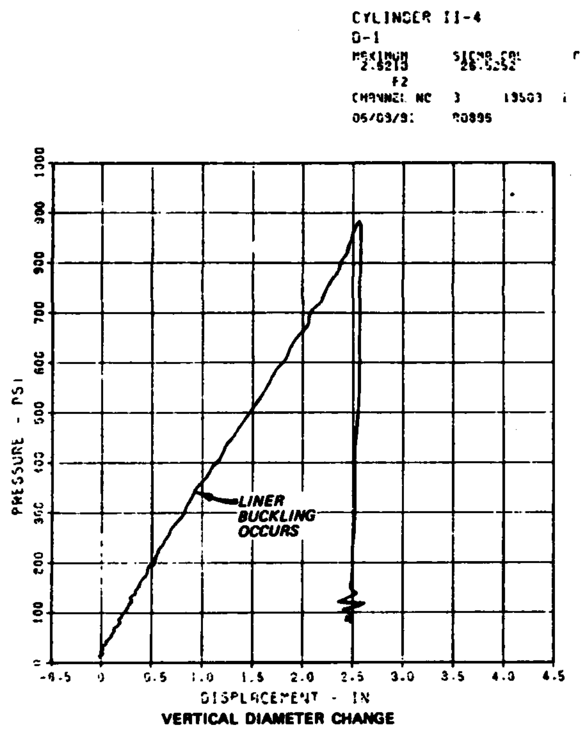


Figure 13. Decrease in vertical diameter vs. surface overpressure for a baseline static test.

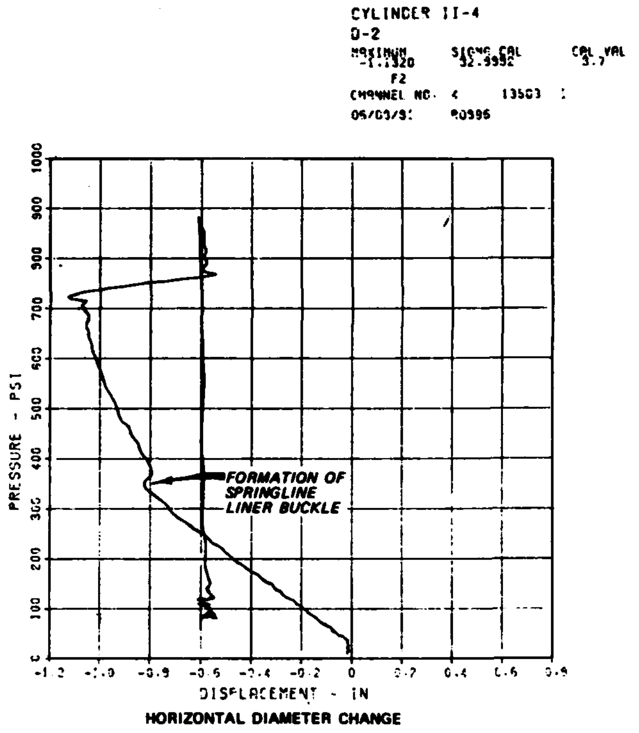


Figure 14. Elongation of horizontal diameter vs. surface overpressure for a baseline static test.

Tests which were carried past the defined failure point showed that springline buckles continued to enlarge with additional cylinder ovaling and concrete crushing, and buckles also appeared as the invert hinge developed. Shown in Figure 15 is an end view of a tested baseline model having failure criteria damage, for which a residual vertical diameter change of 1.3 inches was recorded. Shown in Figure 16 is a baseline model tested beyond failure criteria to the point of invert hinge buckling. This model had a residual vertical diameter change of 2.1 inches.

Effects of Investigated Variables

Most of the Phase I shelter models were tested at a shallow 3-inch burial depth, and were less consistent in the type and location of failure mechanisms. Failures often occurred at the invert with loading curves showing slightly less response stiffness than those of similar Phase II tests. Shallow burial depths seemed to place tests near a breakpoint in response mode, as similar tests were sometimes observed to differ in deformed shape and failure type.

Variation of wall thickness in the Phase II tests did not affect response mode (Figure 17). Models having thicker walls showed a trend toward slightly stiffer response. Failure overpressures of the wall thickness effects tests were all within the range of those for baseline tests, although ovaling was less severe at the time of failure for models having thicker walls.

No significant response mode or failure point departures were experienced when the inspection panel was deleted nor when stiffness of soil backfill was increased to that of the in situ. However, the structure with no inspection panel did respond slightly stiffer than baseline models.

The explosive environments for the four dynamic baseline tests simulated weapon yields of approximately 0.5 to 1.0 kilotons, and ranged in peak simulation overpressure from about 1000 to 3000 psi. In the most severe of these environments, structural damage was just reaching failure criteria condition, with liner buckles visible just below springlines (Figure 18). This model had a residual vertical diameter change of 2.0 inches.

CONCLUSIONS

1. The baseline cylinders responded statically in an ovaling mode which caused failure mechanisms to occur at the springlines in the form of liner buckles due to severe crushing of concrete, followed by buckling at the invert.

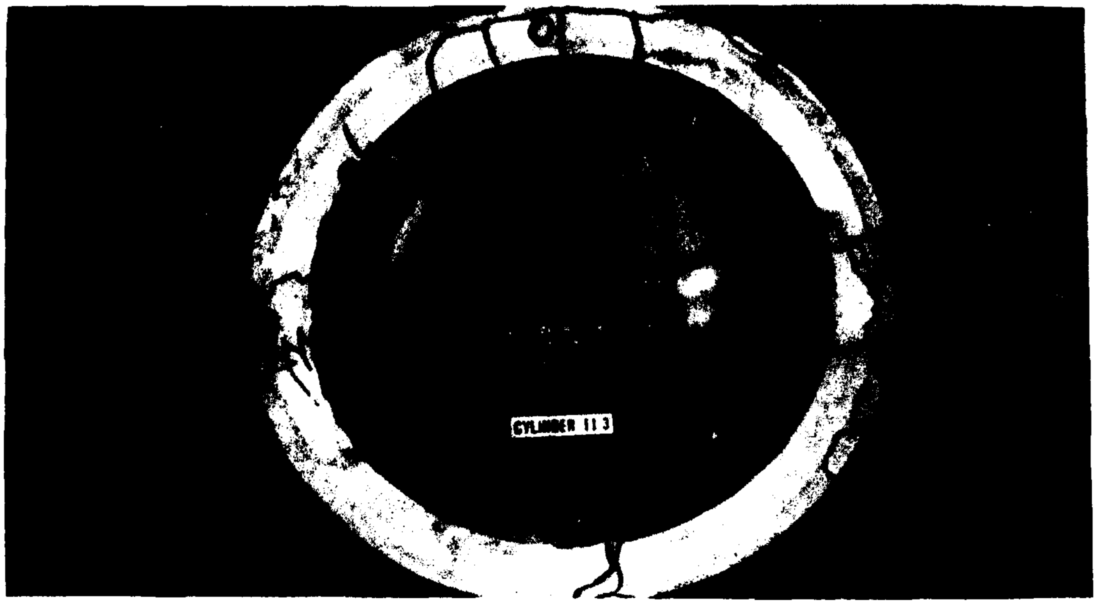

CYLINDER 11 3

Figure 15. Baseline model at failure criteria.

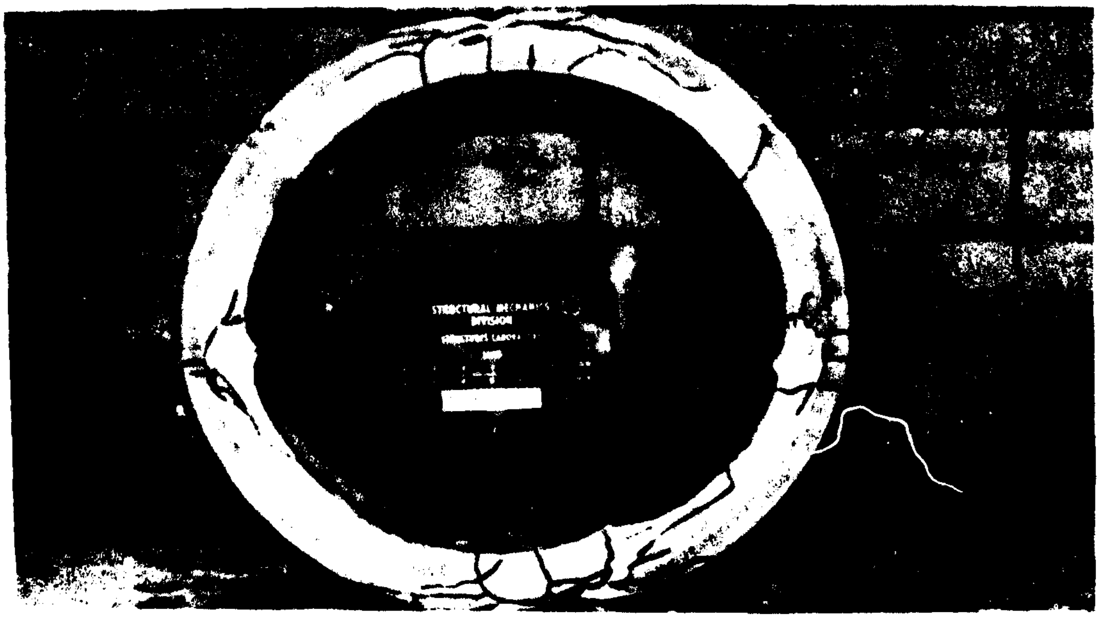
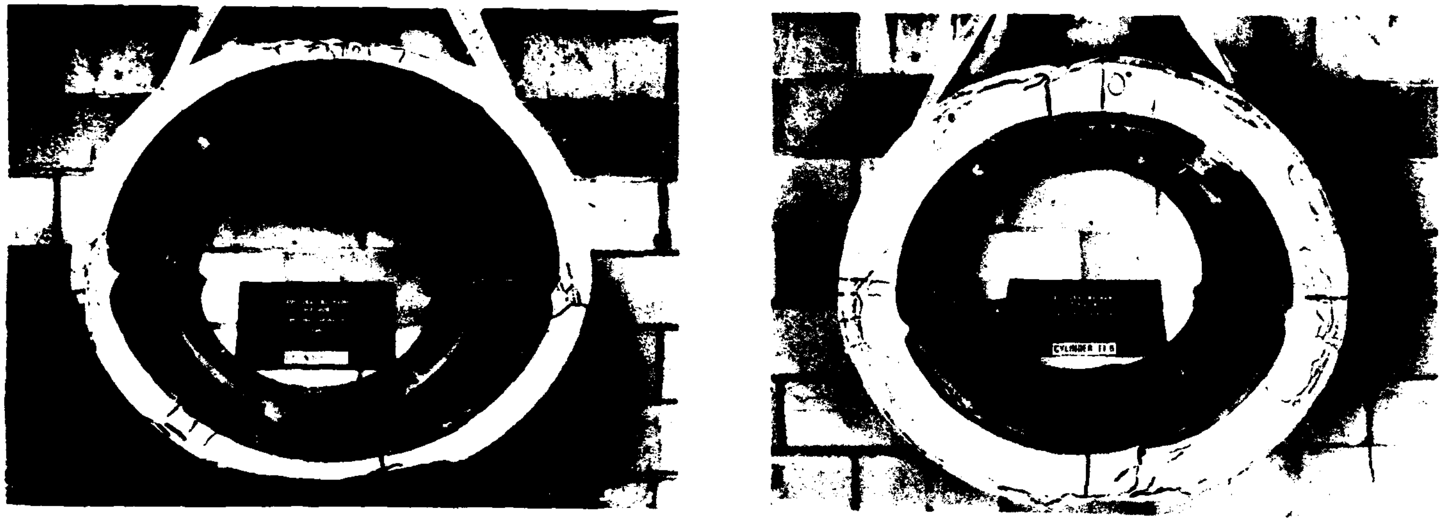

STRUCTURAL MECHANICS
DIVISION

Figure 16. Baseline model loaded past failure to invert buckling.



a. 1.8-inch thickness.

b. 3.6-inch thickness.

Figure 17. Effects of wall thickness.

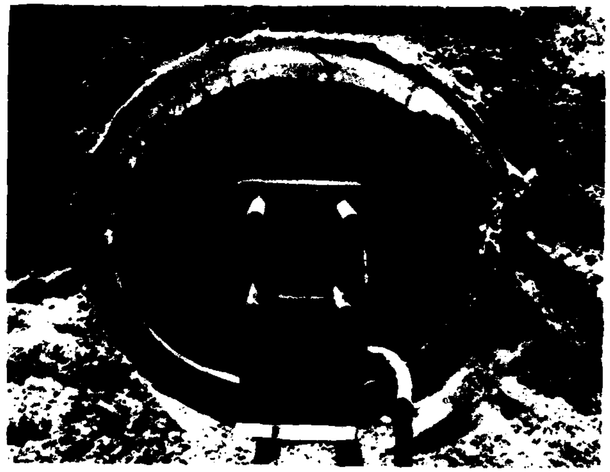
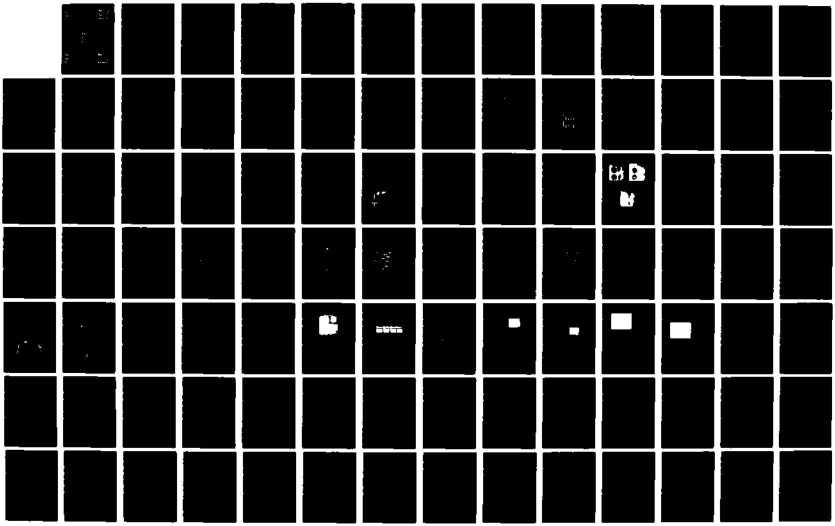
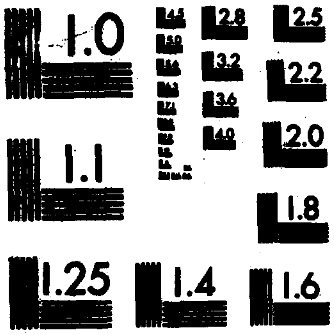


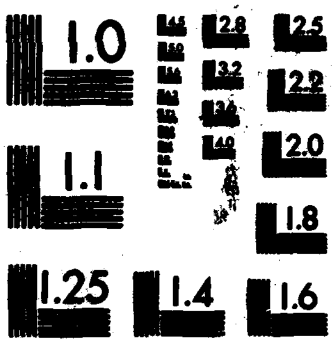
Figure 18. Dynamic test response.

AD-A120 811 PROCEEDINGS OF THE 1982 ARMY SCIENCE CONFERENCE HELD AT 4/6
THE UNITED STATES..(U) DEPUTY CHIEF OF STAFF FOR
RESEARCH DEVELOPMENT AND ACQUISITIO.. 18 JUN 82
UNCLASSIFIED F/G 5/2 NL

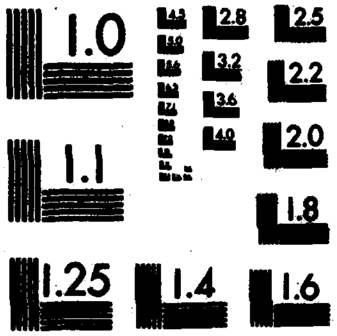



1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

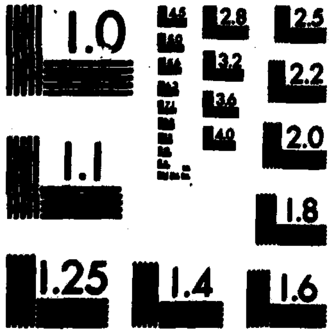
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

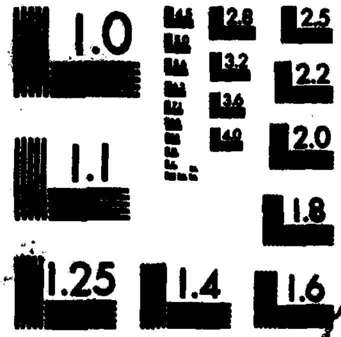
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

2. Tests indicated that the baseline cylinders resisted static surface overpressures of at least 350 psi before reaching a damage level that was compatible with the defined failure criteria (initial buckling of the liner).

3. Cylinders buried at deeper depths responded in a stiffer mode (less bending) and the observed failure mechanism was very consistent for all tests at the deeper burial depths.

4. Cylinders with thicker walls tended to respond more stiffly; however, the basic response mode did not change.

5. The cylinder with no inspection panel and the cylinder backfilled with stiffer soil responded about the same as the other cylinders tested.

6. Dynamic tests of the baseline cylinder confirmed that the dynamic response was similar to that of cylinders tested statically.

7. For a 1.0-kiloton weapon simulation, the ratio of dynamic to static overpressures causing a damage level of the defined failure criteria was found to be in the range of 5 to 10.

MODELING WEAPONS CREW PERFORMANCE

* LLOYD M. CRUMLEY, Ph.D.
ROBERT C. SCHWALM
U. S. ARMY RESEARCH INSTITUTE
FORT SILL FIELD UNIT
FORT SILL, OKLAHOMA 73503

The U. S. Army Research Institute for the Behavioral and Social Sciences is currently conducting a research program to develop a method and obtain the data required to evaluate the ability of weapons crews to maintain their performance over the course of extended, continuous operations. The basic approach involves developing a computer-based model which will meld modifications of selected industrial "methods engineering" techniques with appropriate human performance data in order to predict crew performance without the need to create, train and measure each crew structure alternative in simulated battle conditions.

The approach involves creating a computer-based model which will simulate a crew performing all, or selected portions of, the tasks required to operate the system being studied. The model operates with inputs representing tasks and task times and, when such data become available, it will contain tables defining the performance decrements which are expected to occur during extended operations. Two portions of the project have been completed to date. A crew simulation model has been developed and a task and task time library has been compiled for M109A1 howitzer sections (1, 2). Progress has also been made on a literature review being conducted to determine what data are available for use in developing performance decrement tables.

In future battles howitzer crews will likely be required to operate around the clock for periods of up to eight days. During these periods crews will move as often as 6 to 12 times a day and fire at rates varying from 125 to 500 rounds per 24 hour period. Figure 1 shows a fire rate by day of battle estimate which has been used in past Field Artillery School presentations of a first battle in Europe scenario. In estimating the number of moves, it was assumed that on the average a howitzer would move once for every 50 rounds fired, as the result of repositioning required either because of changes in the location of the Forward Edge of the Battle Area (FEBA) or because of relocations to avoid counterfire. In the projected scenario, the section will fire 2060 rounds and move 41 times during

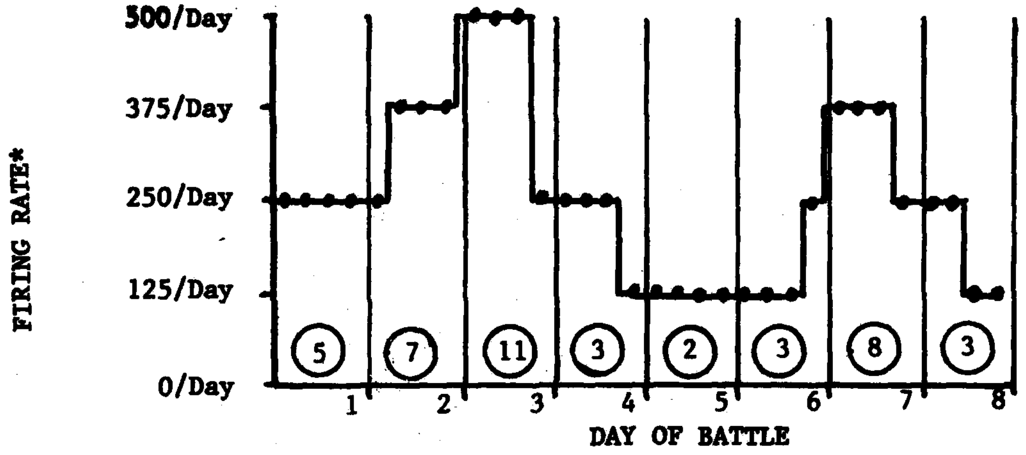


Figure 1. Estimated firing rate per six hour period and moves per day for a potential first battle in Europe scenario.

*In rounds per howitzer per day.

(XX) Moves per day.

the eight day battle. Can a ten man howitzer crew do this and if so, how should the crew workload be assigned in order to best accomplish it? These are the questions that the research program, reported here in part, seeks to answer.

ASSUMPTIONS UNDERLYING THE MODEL

In order to model crew structures and analyze a crew's ability to withstand the demands of the extended, continuous battle we found it necessary to address first the various kinds of tasks which a howitzer crew must perform. We considered three parameters: the type of task, the purpose of the task and the impact of the task on the section's resources. The classification scheme describes tasks as discrete or level-of-effort, fighting or support, and expending or replenishing. Replenishment tasks either replenish resources or reduce risks.

A discrete task is one that has a start and [illegible] oint and is limited by the task itself in terms of the number of perso [illegible] an work on it. An example is selecting a fuze and putting it on the projectile. Such tasks are variable in the time it takes to perform them and when the task is over the crew member is free to go on to another task. If discrete tasks are performed slowly, or if task sequences are performed slowly because of poor task sequence arrangement, the affected crew members are not free to go on to other tasks. Discrete tasks can be performed more rapidly with effective training and they lend themselves to rearrangement to improve the crew effectiveness. Level-of-effort (LOE) is applied to all tasks which are not considered as discrete. This category includes such activities as standing guard, preparing defensive positions and sleeping.

The reason for the discrete vs level-of-effort task distinction is to simplify the model. The tasks we have classified as discrete lend themselves to analysis by the modeling approach we have adopted. Discrete tasks can be assigned in different orders and to different crew members; they can be evaluated individually to determine the effects of various parameters, such as training or fatigue; and they are all involved in crew processes in which the quicker they are properly performed the better the crew performs. LOE tasks are a scheduling problem rather than a crew structure problem. LOE tasks tend to take longer to perform, do not lend themselves to a "quicker is better, so we should train to perform them faster" approach and usually have little or no room for being speeded up by a more efficient method.

In developing a rationale for our model we have also considered which of two purposes a task serves. Fighting tasks include all those tasks which must be performed when a howitzer section is emplaced, fired, or march-ordered (e.g., loading a projectile). Support tasks are all the tasks not classified as fighting tasks. Support tasks are performed in order to maintain the ability of the section to continue to perform its fighting tasks (e.g., ammunition resupply).

Finally, tasks may be considered in terms of their impact on the section's resources. The fighting tasks all expend resources, e.g., firing uses ammunition, moving uses fuel and men get tired. Support tasks, on the other hand, are intended to restore or conserve resources; hence, they either replenish resources expended as the section moves and fires, or they reduce the risk that the section will be placed in a position where its resources will be unnecessarily expended.

The task categories provide a basis for considering the tasks of a howitzer crew in a series of steps rather than in a single stage. Thus, in modeling a crew, and in actual operations, deciding if a crew can perform adequately can be answered by determining first if they fight the howitzer adequately and then if they have time to perform the necessary support tasks. Once this baseline has been established, performance during extended, continuous operations can then be considered.

Three of the four task sequences involved in fighting a howitzer--emplacement, firing, and march-order--are composed almost entirely of discrete tasks. The fourth sequence, the in-transit phase, is composed mostly of LOE tasks. It is feasible, therefore, to model the set-up, firing and march-order sequences and evaluate how well different crew arrangements meet time standards for performing these phases. This approach makes speed of performance a measure of a crew's ability to perform the task sequences involved in fighting the section. (Errors are taken into account by their effects on speed. When errors are made, they must be corrected before proceeding, hence speed suffers.)

The number of persons not assigned to fighting duties determines the amount of time available for performing support functions. Judgments must be made concerning the priority of the various replenishment and risk reduction tasks in order to assign the available personnel most efficiently. How well a crew structure meets the support task requirements is then measured by matching the support task man-hour requirements to the man-hours the crew structure provides to attend to them.

A crew structure is adequate when it meets the time standards for fighting task sequences and provides enough man-hours for the required support tasks. Structures that do not fully meet these requirements become increasingly inadequate as a function of both the lack of adequate support man-hours and the additional time required to perform the fighting tasks. Selection between sub-optimal alternatives for actual operations is a problem that requires the kinds of data that the Crew Performance Model can provide.

The present standard crew operating procedure is defined in Technical Manual 9-2350-217-10N (3). The procedure indicates that ten men emplace a howitzer section, eight fire it, nine march-order it and either nine or ten men are in-transit as the section moves between locations. This operating procedure provides little manpower for assignment to support duties. For instance, when the gun is firing, only two persons are available for support duties. Thus, if a section were to remain in one location for an entire day, 48 man-hours (2 X 24) would be available for support duties. Since four hours sleep per man each day is a minimal support function, 40 of these 48 hours would be required just to provide the crew with minimal rest.

Table 1 is a list of support tasks classified as replenishment or risk reducing. The task times were obtained mostly from doctrine requirements and Field Artillery School documents. Table 1 shows that support requirements easily exceed 129 man-hours per day. Thus, the support tasks which must be performed if a howitzer section is operated around the clock and according to doctrine will almost certainly exceed the ability of the crew to perform them. For that reason Section Chiefs continually face a series of trade-offs: They must first decide how many people can be removed from the fighting tasks; they must then evaluate the situation and set priorities on the replenishment and risk reduction duties required to support the section and battery; finally they must schedule the available support manpower to perform the most essential support tasks.

Table 1

List of Replenishment and Risk Reducing Support Tasks and Selected Estimates of Performance Time

		Estimated Time in Man-Hours/Day
	Replenishment Tasks	
a.	Replenish ammunition (500 rds)	16.0
b.	Replenish POL	1.0
c.	Non-scheduled maintenance	4.0
d.	Sleep	40.0
e.	Supply SGT duty	6.0
f.	Personal maintenance (chow, hygiene, body functions, weapons and gear)	4.0
g.	Identify and prepare new firing positions	12.0
	Risk Reduction Tasks	
a.	Preventative maintenance	1.0
b.	Perimeter defense/early warning	20.0[1]
c.	Guard nuclear ammunition	5.0[2]
d.	Prepare supplementary positions	N/A
e.	Prepare latrines, sleep areas, etc.	N/A
f.	Improve positions	N/A
g.	Camouflage	N/A
h.	Screens	N/A
i.	Foxholes	N/A
j.	Crew-served weapon positions	N/A
	Involuntary Downtime while Moving	20.0
	Total of Estimates	129.0

[1] This assumes two guards at each of four stations with the battery operating as an eight-gun unit. If the battery were operating as two four-gun units, this figure would be 40.0.

[2] This would be 10 man-hours/day if the battery were operating as two four-gun units and both units had nuclear ammunition.

NOTE: N/A indicates time estimates for these tasks were not available.

STRUCTURE OF THE CREW PERFORMANCE MODEL

The ARI Crew Performance Model was developed to assist the Section Chief--or the system developer if the system is not yet fielded--in answering such questions as how to structure a crew to deal effectively with both fighting and support functions under a variety of conditions. The modeling approach we use takes into account the tasks to be performed, the order of performance, the time it takes to perform each task, and the number of men doing those tasks. The model consists of a task library, an input program, and the software to handle that information.

The task library contains the tasks and subtasks which must be performed by the members of an M109A1 howitzer section. (A task library could be built for any system, provided the discrete tasks making up that system can be defined.) Each task is narrowly defined, has minimum, average, and maximum completion times, and takes into account any requisite or concurrent tasks. Each library entry also has a task type code, a probability of performance value and an indication of how the time data were obtained. The format for task library entries is shown in Figure 2, Part A. Entries were obtained from appropriate documentation, discussions with subject matter experts and videotapes of two III Corps howitzer crews in operation. For the latter, each crew emplaced their section, fired three rounds and march-ordered; this cycle was performed four times. Thus, our times are generally based on eight trials for emplacement and march-order tasks and 24 trials for firing tasks.

In order to use the model, the user--whether Section Chief or system developer--needs to specify the number of persons to which tasks will be assigned and then assign the tasks in the order each crew member will perform them. An example of an input sequence is shown in Part B of Figure 2. The main program then simulates the number of iterations specified and provides the output shown in Part C of Figure 2.

USING THE MODEL

The model has been used to determine the effects of various crew sizes/task assignments on the performance of howitzer sections. The base case is the crew size/task structure as presented in the TM. For variations of those structures, we deleted men and assigned their tasks to the remaining men in such a way that those crewmen who were not very busy tended to assume most of the tasks, particularly those which were logically consistent with their current positions. The results of several runs are shown in Figure 3. The figure shows completion times and percentages of crew member idle time for units of different sizes performing fighting task sequences. In the figure, completion time is expressed as a percentage of the median time the model calculated for a full ten man crew performing the task sequences according to the Technical Manual.

```
A                        TASK LIBRARY ENTRY

   123  456  789  2.0  3.0  4.5 TT.PPP  S
      AN ILLUSTRATION OF A TASK SHOWING ALL VARIABLES

B                     CREW TASK/STRUCTURE INPUT

   1  400  402  404  406  408  410  412  650  652  654
   2  555  556  557
   3  558  559
   4  502  504  507  508  510  512  600  602  516
   5  454  452  456  468  472  450  458  465  471  473

C                           MODEL OUTPUT

   COMPLETION TIMES, FASTEST TO SLOWEST, EVERY 10TH PERCENTILE
134.0161  152.3820  161.5413  166.9992  172.9557  177.7338  183.3907
188.7629  197.0742  203.3185  230.5410

CRITICAL MAN   1    0      2    0    3    0     4     0      5  200

IDLE TIME      1 75.6      2 90.4    3 91.0     4  66.1      5 18.0
```

Figure 2. Examples of task library entries, model input, and model output.

Figure 3 shows that reducing from ten to six the number of men assigned to set-up and boresight has only a small effect on completion time. This is because the improved efficiency of assigned crew members generally compensates for the lesser number of men. The six man unit is on the average only 5% slower than a ten man unit, but as unit size is further reduced to five the time penalty increases to 11.0%. Neither increase in completion time appears substantial given the corresponding increased availability of support personnel.

Data in Figure 3 also indicate that a reduction from eight to five men during firing adds a first-round firing time penalty of 8.6% per mission but provides five men to perform support duties. Further reducing the fighting unit to four would cause the time required to fire a single-round mission to increase by 12.7% over base rate.

The TM sequence for march-ordering a howitzer section involves nine men since the Gun Guide is usually elsewhere establishing a new location. Here again a reduction in crew members makes very little difference until a five man unit is used. Figure 4 shows that five man units accrue a 14.7% time penalty, but the penalty increases to 28.6% when only four men perform

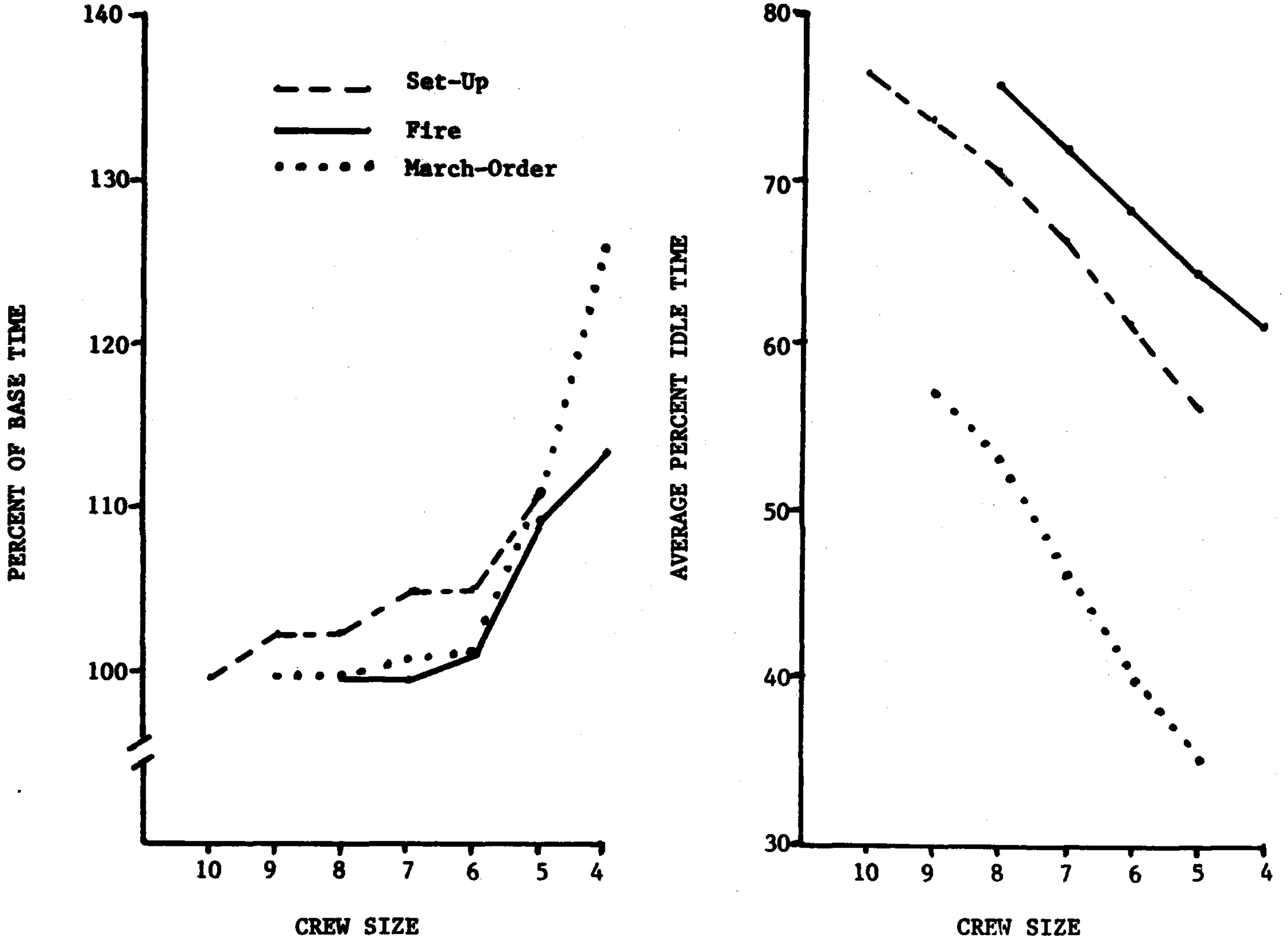


Figure 3. Times required to set-up, fire and march-order an M109A1 howitzer section and crew member idle times for various crew sizes.

the march-order tasks. Thus, four men could be made available for support duties without significantly impairing performance on march-order tasks.

From this analysis of crew structure effects on set-up, firing, and march-order times, it appears that fighting tasks would not necessarily be impaired when operating with a reduced crew if the crewmen are more efficiently assigned their tasks. By doing this, not only is idle time reduced but personnel are freed to perform necessary support duties.

GOALS FOR WEAPONS CREW STRUCTURES

In considering the results of our simulation of howitzer crew operations, we have formulated five goals and have examined our data to determine which crew structure best meet these goals. The goals we think must be met by an adequate crew structure are:

1. Be able to emplace, fire, march-order and transfer the section in an acceptable time or at an acceptable rate.
2. Permit enough people to be assigned to the support functions that lack of support does not reduce the section's fighting ability to an inadequate level.
3. Have a reasonable capability to respond, in a degraded mode, when section personnel are lost.
4. Be effectively applied and practiced in a peacetime milieu.
5. Transfer to battle conditions with a minimal burden on the crew.

Data generated by the ARI Crew Performance Model imply that howitzer crews can be significantly modified to provide a better base for fighting a howitzer section. There are several ways in which the existing crew structure can be changed. One particularly attractive rearrangement is to divide the ten man crew into two virtually identical five man units (5/5), a split-crew approach, that alternates the units between the fighting and support tasks.

A 5/5 split-crew structure is proposed in Table 2 where it is compared to the crew as it is designated in the Technical Manual. The TM crew structure consists of ten positions, and one person is assigned to each position. The 5/5 crew structure provides two units, a "Red Unit" and a "Gold Unit", which are identical except that one unit contains the Section Chief and the other is supervised by a Unit Chief. During extended continuous operations, these units would alternate between the fighting tasks and the support tasks. A summary of the fighting task decrements and the time available for support tasks for such a division is shown in Table 3.

When a gun is emplaced according to the sequence defined in the TM, the entire ten man crew, including the Gun Guide, is involved. In the 5/5 solution, five firing unit members and the Gun Guide (6 men) are involved. The median set-up time is only 5% longer for the 5/5 crew than the standard

Table 2

TM and Split-Crew Structures

TM STRUCTURE

Chief of Section	(CS)	Cannoneer #1	(#1)
Gunner	(G)	Cannoneer #2	(#2)
Assistant Gunner	(AG)	Cannoneer #3	(#3)
Howitzer Driver	(HD)	Cannoneer #4	(#4)
Section Driver	(SD)	Cannoneer #5	(#5)

SPLIT-CREW STRUCTURE

Red Unit		Gold Unit	
Chief of Section	(CS)	Unit Chief	(UC)
Gunner	(G)	Gunner	(G)
Cannoneer #1	(#1)	Cannoneer #1	(#1)
Howitzer Driver	(HD)	Howitzer Driver	(HD)
Section Driver	(SD)	Section Driver	(SD)

crew. This is a rather minor time penalty and, in a real sense, even this may not constitute a liability. Note that with the 5/5 structure, four crewmen are available during the emplacement process to perform support functions. In a hostile environment these four men would most logically be assigned to set up camouflage or emplace screens. Therefore, when the gun is reported up and ready in the 5/5 case, some of the support work would already have been completed; in the TM case the support work would all remain to be done. Thus, the actual time required to have the gun both up and ready and protected would probably be less in the 5/5 case.

The effect of shifting to a split-crew structure is somewhat greater for firing than for emplacement. We estimate the median time to prepare and fire a first round increases by 8.6% if the 5/5 structure is used. Again, however, this is compensated for by there being five rather than two men available to perform support functions.

A more subtle form of compensation during the firing cycle (or any other sequence) is that with only five men involved, there are no excess crew members. Consequently, there is a real pressure to improve crew effectiveness by removing problems rather than merely working around them by applying another pair of hands. Indeed, when the work flow during firing was examined we identified at least three materiel changes which, if implemented, would make the five man firing operation faster than the standard eight man operation. For example, because howitzer crews are not issued strap cutters and must improvise accordingly, cutting straps averages 41 seconds per eight-round pallet. Therefore, an average of about five seconds per round fired is the cost of the poor equipment furnished to cut straps. Clearly, if a strap cutter were placed in the tool kit on the section vehicle, the time required to open a new pallet would be reduced considerably--probably to under five seconds--with a consequent reduction of median one round firing time by about four seconds. When the model was used to evaluate the effect of a hypothetical situation in which this and other materiel changes were included, we found that the median completion time for the 5/5 structure during firing was actually faster than the present eight man approach.

Table 3

Times to Complete Fighting Task Sequences and Times Available for Support Duties for the TM and Split-Crew Structures

Activity	Men Available For Fighting Functions		Men Available For Support Functions		5/5 Time Penalty
	TM	5/5	TM	5/5	
Emplace	9 + GG	5 + GG	0	4	5.0%
Fire	8	5	2	5	8.6%
March-order	9 + GG	5	0	5	17.7%
In-transit	9 + GG*	9 + GG*	1	1	0.0%

*The Gun Guide is normally not in-transit with the crew.

The march-order sequences constructed for five man units also shows a time penalty, adding 14.7% to the nine man median completion time. However, this decrement is more illusory than real. The march-order base case follows the TM assumption that nine men are immediately available so that the march-order sequence begins with all nine men working. In an actual battle situation, however, a hasty departure would surely begin with some of those nine men assigned to perimeter defense or other support functions. Thus, nine men are not immediately available for work. The apparent 14.7% decrement would have been less if the TM sequence did not begin unrealistically with eight men at the gun and one nearby (the Gun Guide having already departed). Actually, given the number and diversity of the support tasks which must be performed, the 5/5 structure with an independent, self-contained support unit may well be faster.

The in-transit situation would be unaffected by the change to a 5/5 crew structure. In the 5/5 system, the unit assigned to the fighting tasks would provide two drivers, two air guards, and an NCO in-charge. The support unit would provide a gun guide who would probably be off with the advance party. The other four support unit members would do those minor personal support duties which are possible while being transported or else they would be idle. This is essentially the work/idle situation that presently exists while in-transit.

It seems that a 5/5 division provides no significant disadvantages in the emplacement and march-order sequences and that the decrement which accrues in the first-round-out situation can be overcome with minimal improvements in the equipment available to a howitzer crew. Thus, a 5/5 crew split meets the first goal--fighting the howitzer in an acceptable time--as well as the present TM crew structure.

The second goal--having enough men available to support the section--is clearly better met by the 5/5 split-crew structure. Indeed the 10, 8,and 9 men specified for the set-up, firing and march-order sequences in the TM structure leave at most only two men to perform support duties. The split-crew structure makes available five support personnel at all times.

The 5/5 structure is also well suited as a point of departure for degraded mode operations (Goal 3). Battery losses of up to one man per section can be absorbed by reducing the support task units from five to four men. If further battery losses occur, men could be pulled from fighting task units until sections are operating in a 4/4 condition. In that case, sections, and the battery to which they are attached, would be operating more slowly and at a greater risk of suffering either adverse actions or having inadequate replenishment capability. More likely, however, the battery with 20% casualties would have lost essential materiel. Personnel from inoperative sections would then be available to other sections as replacements to create five man firing units; fewer viable sections would, of course, be the result.

Differences between the TM assignment method and the 5/5 concept also favor the 5/5 concept when peacetime manning problems are considered (Goal 4). Two basic manning problems exist in the present Field Artillery cannon batteries: Many batteries do not have a full complement of personnel, and high levels of turbulence characterize most section crews. These two problems have far more significant consequences for the TM structure than they would have for the 5/5 structure. The most obvious problem associated with using the 10/8/9 set-up/fire/march-order structure is that a section must have ten crew members to use it. Realistically, this is seldom the case; hence, there is usually no opportunity to create or to practice ten man crews. A 5/5 split-crew structure, on the other hand, requires fewer personnel in order to form units for training or fighting. The TM structure is also particularly vulnerable to the effects of turbulence. The absence or shifting of personnel from a crew because of internally created turbulence or the replacement of existing crew members by new personnel creates situations that are hard to overcome since section activities are not currently standardized. As crews now exist there is little likelihood that a crewman serving in a particular position on a particular crew will perform the same tasks when serving the same position but on another crew. In contrast, with a 5/5 split-crew structure, task assignments would be standardized, thereby minimizing the need to retrain when personnel change positions or crews.

As for Goal 5, the 5/5 structure appears superior for moving from a peacetime to a wartime situation with minimal impact. Standardization, which can be readily accomplished with the 5/5 approach, makes it possible to bring a battery up to full strength by merely adding five man units. As new units arrive at a battery, Section Chiefs simply convert sections from half manned sections to fully manned sections. This saves Section Chiefs from the need to move from a poorly defined crew structure to the ten man TM structure, a structure which as we have suggested is not very efficient.

PERFORMANCE DURING CONTINUOUS OPERATIONS

Our discussion until now has shown how the ARI Crew Performance Model can estimate crew size/structure effects under first-day or near optimal conditions. In a more realistic scenario, however, howitzer crews will likely be fighting for extended periods--up to 8 days--under high work/low rest battle conditions. Anecdotal and experimental evidence on the effects of these conditions on performance suggests that the effects will be both negative and substantial and will increase as the duration of combat and sleep loss increases. For that reason the model has been equipped to adjust first-day task performance as some function of accrued impairment from Day 1 through Day N. Also, because different types of tasks will be differentially affected by sleep loss, the model is equipped to apply decrements according to task type.

In order to determine what data are available for estimating long-term performance under the conditions noted above, we have conducted an extensive literature search and review of the effects on performance of total and partial sleep loss and work/rest cycles. Only research on behavioral effects, vis-a-vis subjective and physiological measures, were considered. Both the open literature and government sources were searched. The results of this review will be reported elsewhere in detail. We also examined various task taxonomies to determine how the howitzer tasks and tasks in the literature might best be classified. In the end, none was suitable, primarily because of the nature of the measures used in the literature, so we prepared a rather global classification scheme for our use.

The literature describing performance after varying periods of total sleep loss provided over 205 titles, with approximately two-thirds of these studies testing performance after less than 72 hours of total sleep loss. Approximately 40 of these articles documented information useful for determining decrement values for use in our model. Considerably less research was found on the effects of partial sleep loss and on performance under relevent work/rest ratios. In fact, only 55 titles in these areas were found and fewer than ten have documented evidence suitable for use in constructing decrement tables for our model.

Table 4

Documented Studies of the Behavioral Effects of Total and Partial Sleep Deprivation as a Function of Task Type and Time

TEST DAYS	1	2	3	4	5	6	7	8
Physical Strength	T	T	T	T				
Gross Body Movement and General Coordination	T	T	T	T				
Fine Movement and Eye-Hand Coordination	T	T	T	T	T P	P	P	P
Following Directions and Short-Term Memory	T P	T P	T P	T P	P	P	P	P
Rules, Regulations and Long-Term Memory	T	T						
Problem Sensitivity and Responsiveness to Change	T P	T P	T P	T P	T P	T P	P	P
Decision Making and Taking Action	T		T					
Mental Effort, Attention and Concentration	T P	T P	T P	T P	T P	P	P	P

T = Total sleep loss. P = Partial sleep loss or work/rest ratio.

Table 4 is a general overview of our survey with the data available from the literature presented as a function of task type and duration of test period. Separate entries are shown for total sleep deprivation and work/rest ratio studies. Very little work has been done in these areas and most of what has been reported deals with periods of only a few days. Of the studies testing performance under realistic conditions, only one tested performance after an initial period of total sleep deprivation followed by periods when four hours of sleep per day--the maximum likely--were allowed. Some task types are notable for their lack of research documentation. Furthermore, within task types, a failure to standardize tasks and procedures contributes to considerable difficulty in interpreting results. Indeed, most of the studies represented in Table 4 suffer from one or more methodological problems of varying degrees of severity. Quite common, for instance, is failure to control for learning or circadian rhythm effects.

Thus, an extensive review of the literature on the effects of extended performance with reduced sleep illustrates the dearth of useful information available. For this reason, only a portion of the decrement table for use with the model can be completed based on data in the available literature, and even that portion, because of problems in the original research, is of limited value.

CONCLUSION

The ARI Crew Performance Model can be used as an analytic tool for evaluating the effects of crew structure on performance. Whether used with developing systems or systems in the field, the model can be of value in determining the effects of changes in crew size and the assignment of tasks on the relative efficiency of varying crew structures. However, before the model can be fully utilized, particularly when crew performance in a realistic battlefield scenario is being considered, a great deal of work remains to be done to determine the effects on human performance of extended, continuous performance under conditions of high work/low rest periods.

REFERENCES

1. Schwalm, R. C., Crumley, L. M., Coke, J. S., & Sachs, S. A. A description of the ARI Crew Performance Model (RR 1324). Fort Sill, OK: U. S. Army Research Institute for the Behavioral and Social Sciences, 1981.

2. Coke, J. S., Crumley, L. M., & Schwalm, R. C. Emplacing, firing and march-ordering an M109A1 howitzer: Tasks and task times (RR 1312). Fort Sill, OK: U. S. Army Research Institute for the Behavioral and Social Sciences, 1981.

3. Department of the Army. Technical manual operation and maintenance manual (User) for howitzer, medium, self-propelled: 155mm, M109A1 (2350-00-485-9622) (TM 9-2350-217-10N). Washington, DC: Headquarters, Department of the Army, 1978.

THE DEVELOPMENT AND EVALUATION OF HEMOGLOBIN SOLUTION AS A BLOOD SUBSTITUTE (U)

*FRANK DEVENUTO, Ph.D.
ANGELO I. ZEGNA

LETTERMAN ARMY INSTITUTE OF RESEARCH
DIVISION OF BLOOD RESEARCH
PRESIDIO OF SAN FRANCISCO, CA 94129

The concept of using hemoglobin (Hb) solution as a resuscitation fluid in blood replacement therapy has been advocated for many years and has led to extensive in vitro and in vivo investigations (1,2). These studies have been pursued for several reasons. One of the reasons is that blood, a scarse and valuable resource, has a limited storage life, which at present does not exceed 35 days and requires specialized expertise and equipment for its collection, transportation, storage and pre-transfusion preparation such as typing and cross-matching. Furthermore, problems have occurred during and after transfusion of blood. These problems include hemolytic transfusion reactions, allergic reactions, circulatory overload, embolism, citrate toxicity, coagulation disturbance and the transmission of disease (3-4). Therefore, a strong incentive exists for finding a suitable resuscitating solution which would alleviate or minimize such problems.

Another reason for the development of an effective blood substitute is the ability to administer the resuscitating solution to the patients directly at the place where the injury causing loss of blood volume has occurred, without requirement for transportation of patients to treatment centers or medical facilities. Thus, immediate medical assistance by restoration of blood volume and oxygen carrying capacity will greatly increase the chances for survival. Furthermore, a stable resuscitating solution can be stockpiled and, therefore, can minimize the logistic problems of supply, storage and transportation which occur when large quantities of fluid replacement are necessary in mass casualties.

Significant advantages can be gained by the development of a resuscitating solution capable of transporting oxygen, maintaining oncotic pressure, and being readily available when massive transfusions are required. Stringent requirements must be met for a resuscitating solution in order to be effective. As a blood substitute, this solution not only must be capable of restoring vital functions, but must not elicit permanent adverse effects when administered to patients.

A solution of Hb has the potential to fulfill the characteristics required for a blood substitute. Many investigators (2,5-9) have stressed several advantages of this solution as compared with other resuscitating fluids or plasma expander: Hb is a component of normal blood, can be prepared from outdated human erythrocytes, does not require typing or cross-matching prior to use, is capable of transporting oxygen (5-7,9), has oncotic activity (10,11), has lower viscosity than blood (12,13), does not cause microaggregates (14,15), and may not induce significant immunologic reaction (8,16-18). Furthermore, Hb is highly soluble in physiological solutions and can be stored for extended periods of time (14,19).

DEVELOPMENT AND EVALUATION OF HEMOGLOBIN SOLUTION

In the preparation of Hb solution, two important requirements must be met: first, the Hb must be free of components of the red cell membrane, and second, the method must not include cumbersome and tedious procedures which would preclude its easy adaptation to large-scale preparation. Methods based on hemolysis of red cells and subsequent centrifugation (2,16) to remove particulate matter, even if followed by micropore filtration, give rise to the possibility of contamination with soluble compounds of the red cell membrane. A rapid, simple, and reproducible procedure for the preparation of Hb solution free of soluble and insoluble stromal contaminants has been developed in our laboratory (20). The procedure is based on the crystallization of the Hb molecule in a high ionic phosphate buffer and repeated washings of the crystals prior to solubilization and sterilization. This crystallization method is easily adaptable to large scale preparations to process high volumes of outdated blood.

To stockpile Hb preparations and thus meet the need for supplies when massive fluid therapy is required, Hb must be stable during long-term storage as a solution or, preferably, in a dry form. Investigations were performed on the stability of Hb solutions prepared by crystallization (20) and maintained in blood bags under sterile conditions at -20 C, 4 C, or 25 C. At several intervals, methemoglobin, P_{50}, n-values, osmolality, oxygen capacity, Na, K, and pH were assayed (14). The data in Figure 1 show that solutions maintained at -20 C demonstrated no alterations in the integrity and function of the Hb molecule after 2 years of storage. Solutions kept at 4 C remained stable for 12 months; after 12 months and especially after 18 months, the refrigeration-stored solutions showed deterioration, as expressed by an increase in methemoglobin content and a decrease in P_{50} (50 percent saturation). Hb stored in a liquid form at room temperature demonstrated alterations in the structure and function within four weeks (14).

In the development of Hb as a blood substitute, an important goal would be to convert the Hb solution to a freeze-dried powder for compact

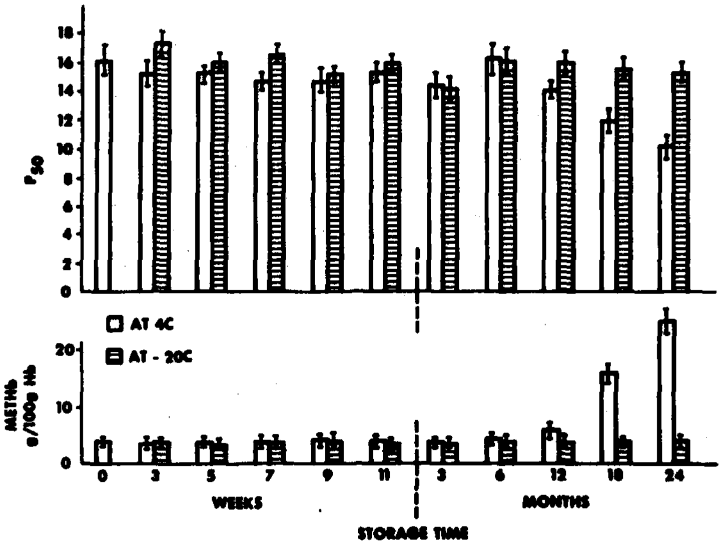


Fig. 1. Methemoglobin content (MetHb) and P_{50} (at pH 7.4) of hemoglobin solutions stored at 4 C or -20 C. Mean ± SD (n=4).

TABLE 1
ANALYSIS OF HEMOGLOBIN SOLUTION LYOPHILIZED IN
THE PRESENCE OF 3 PERCENT GLUCOSE

Assays[a]	Pre-lyophilization	Post-lyophilization and reconstitution
Hb, gm/100 ml	7.7 ± 0.3	7.4 ± 0.2
MetHb, gm/100 gm Hb	3.15 ± 0.61	3.89 ± 0.82
Na, mEq/l	110 ± 4	104 ± 5
K, MEq/l	4.0 ± 0.1	3.9 ± 0.2
Osmolality, mOs/kgm	309 ± 8	294 ± 6
pH	7.31 ± 0.04	7.41 ± 0.04
P_{50}, at pH 7.4	14.27 ± 0.7	14.19 ± 0.9
n value, Hill slope	2.44 ± 0.2	2.36 ± 0.1
O_2, volume percent	10.36 ± 0.63	10.42 ± 0.41
Glucose, gm/dl	2.93 ± 0.10	2.89 ± 0.08
Sterility	Sterile	Sterile

[a] Mean ± S.D., n = 3
Hb, Hemoglobin; MetHb, methemoglobin; Na, sodium, K, potassium, and O_2, oxygen.

storage at room temperature. At the time of transfusion, addition of sterile water to this dry, powdered Hb would yield a solution ready for infusion. Crystalline Hb solution has been lyophilized in our laboratory (19) following addition of compounds to establish protective conditions for obtaining freeze-dried Hb chemically and functionally unaltered. Glucose or sucrose, added to a final concentration of 3%, was most active in protecting the Hb molecule from deterioration (Table 1). Stability studies (Figure 2) demonstrated that Hb lyophilized in the presence of 3% glucose and maintained at 4 C did not show any significant alteration in structure and function for several months. Freeze-dried Hb samples stored at room temperature were unchanged for 6 months but, after this time, a progressive increase in methemoglobin content and decrease in P_{50} were observed.

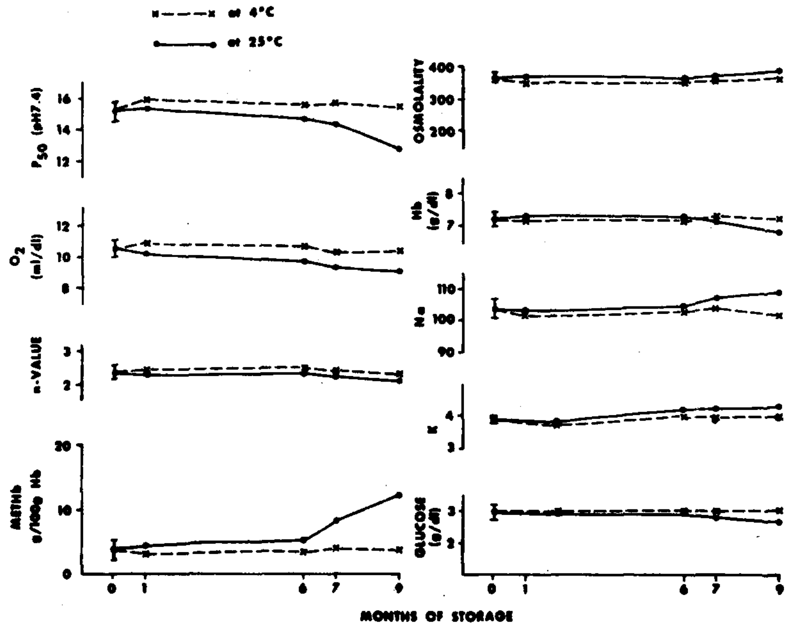


Fig. 2. Methemoglobin content (MetHb); n-value (Hill's slope); oxygen capacity; P_{50} at pH 7.4; glucose; potassium (K); sodium (Na); hemoglobin (Hb); and osmolality of lyophilized hemoglobin stored at 4 C or 25 C. Mean values are plotted. Representative standard deviation (n=3) for each assay is shown for values at time 0.

One important characteristic of Hb solution is represented by its viscosity which is lower than that of blood. This characteristic could be of significant advantage in the treatment of hemorrhagic shock where constriction of the capillary vessels in the microcirculation would dictate the use of a fluid with a lower viscosity than blood for normovolemic hemodilutions. The *in vitro* viscosity of human blood at different hemodilutions with Hb solution has been studied (13) and the results are shown in

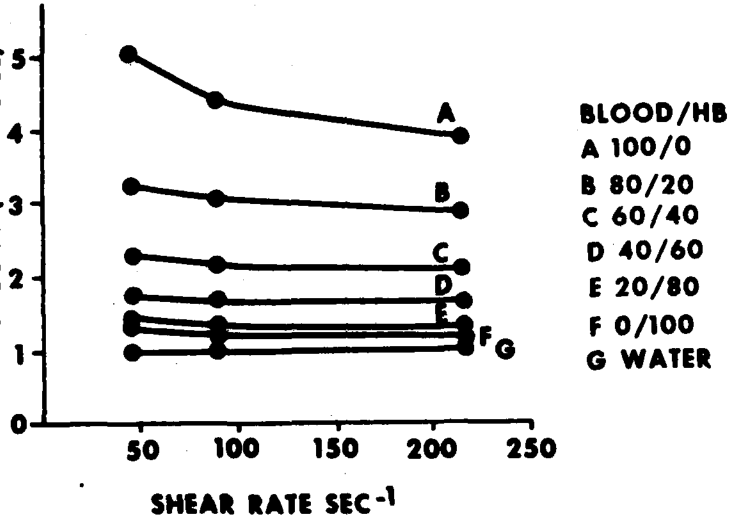


Fig. 3. Viscosity-shear rate relationship for human blood, hemoglobin solution (Hb) and different mixtures of the two components. The relationship for water is also illustrated.

Figure 3 where viscosity is expressed as a function of shear rate. With increasing hemodilution, a decrease in the viscosity is observed at the various shear rates. This decrease is quite pronounced at low or moderate hemodilutions. Hb solution by itself shows a very slight dependence on the shear rate and only at low shear rates. The data demonstrate that hemodilution of blood with Hb solution reduces the viscosity thus improving the blood flow and consequently increasing the supply of oxygen to the tissues through the microcirculation.

IN VIVO EVALUATION OF HEMOGLOBIN SOLUTION

In vivo evaluation of the Hb solution, as prepared in our laboratory, has been done in various animal models transfused with Hb solution to different levels of blood replacement. Exchange tranfusions have been carried out in rats to a total blood replacement of 91 to 93 percent with hemoglobin or albumin solutions (9). The results (Table 2) show that when albumin was used, all animals died at approximately ten minutes after transfusion was completed, whereas all animals transfused with Hb survived for five hours and displayed normal activity during this time. At the end of transfusion, the oxygen capacity of the Hb-transfused animals is about seven times greater than that observed in the albumin-transfused group (Table 2). An explanation for the inability of the Hb-transfused animals to survive longer than five hours was obtained from the results of the plasma Hb disappearance. In these studies, the plasma half-disappearance time of Hb was 3.5 hours, which indicates that at five hours after

TABLE 2
BLOOD EXCHANGE IN THE RAT TO 91-93 BLOOD REPLACEMENT
WITH HEMOGLOBIN (Hb) OR ALBUMIN (Alb) SOLUTION

	SURVIVAL	O_2 (Vol %)
Hb (Ht 3%) (n = 8)	5 hr	9.5-10.4
Alb (Ht 4%) (n = 8)	10 min	1.4-1.8

transfusion the Hb remaining in the plasma represents about 30 percent of the value at the end of transfusion (9). This low Hb level (1.8 g/dl) is not sufficient to maintain vital functions. This disadvantage could be overcome by restoring plasma Hb with repeated bolus injections or with continuous infusion of Hb solution after total blood replacement with Hb solution (21).

Since Hb demonstrated a relatively short vascular retention time, it was important to determine if effective colloidal osmotic pressure (oncotic pressure) was maintained during and after transfusion. Groups of rats were transfused with 7 percent Hb or 7 percent albumin solutions to 25 percent of the initial packed cell volume (11). At this level, which represents

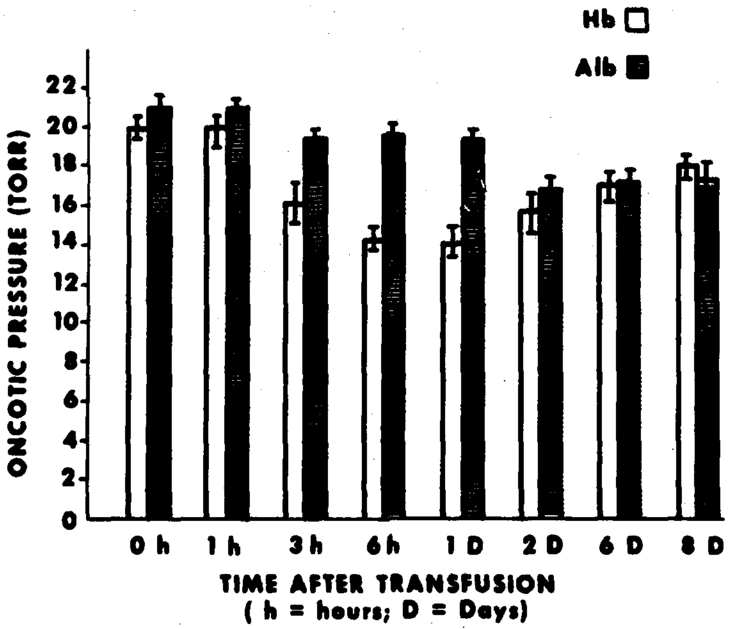


Fig. 4. Oncotic pressure of rat plasma at several times after transfusion to 75 percent blood exchange with hemoglobin (Hb) or albumin (Alb) solution. Mean values ± SD (n=4 rats for Hb and 4 rats for Alb groups at each time).

75 percent blood replacement, all animals survived (9). As shown in Figure 4, immediately after transfusion, both groups showed a similar plasma oncotic pressure of 20 to 21 torr, higher than that observed in normal rat plasma which is 17 to 18 torr at serum protein concentrations of 5.9 percent. During the first 24 hours after transfusion, oncotic pressure values of 21 to 19 torr were determined in the albumin rats, and values of 20 to 14 torr were obtained in the Hb group, reflecting the different intravascular retention time of albumin and Hb. After six to eight days, the oncotic pressure returned to normal values, 17 to 18 torr. Plasma protein electropheresis in the two groups of animals demonstrated a rapid replacement of proteins. By six hours, the intravascular albumin of the Hb-treated group had been replenished to 57 percent and, by 24 hours, to 83 percent of the pre-infusion concentration (11).

In other investigations we have evaluated the disposition and organ distribution of Hb, oxygen transport of human blood hemodiluted with Hb solution, morphologic effects on liver and kidney cells after massive transfusions with Hb solution and several other physiological, hematological and biochemical aspects. The results of these studies have been reported in scientific journals (2,9,13,14,19,20-28).

With increasing knowledge and better understanding of the potential of Hb solution as a blood substitute a clearer picture has been achieved concerning the capabilities and the limitations of the product currently available. This clearer picture has promoted the development of manipulations necessary to modify the hemoglobin molecule for systematic improvements.

RECENT DEVELOPMENTS

Although the oxygen binding values of intra- and extra-erythrocytic hemoglobins are the same, about 1.34 ml oxygen per gram Hb, their oxygen affinities, however, are quite different with P_{50} values of 26-27 and 12-16 mm Hg, respectively. Several factors are responsible for this difference in oxygen affinity, namely the lower Hb concentration as compared to that of Hb in the red cell, the difference between plasma and intracellular pH which can be as high as 0.2 pH units and therefore cause a Bohr effect influencing the oxygen dissociation curve, the qualitive and quantitative change of ions inside and outside the red cell, and the difference in diphosphoglycerate concentration. In experiments in which laboratory animals have been transfused with Hb solution to hematocrit of six percent or lower, it has been clearly demonstrated that the hemoglobin-oxygen complex is capable of satisfying the oxygen demand of the tissues (5-7,9,16,29,30). However, if Hb solution is used as a blood substitute in fluid replacement therapy which does not require extremely large volume replacements and, therefore, causes moderate or low hemodilutions (25-35 percent hematocrit), there is some question as to the ability of Hb to deliver oxygen to the tissues. The low P_{50} may represent a limitation in the use of the Hb solution as presently formulated.

Several studies have shown that plasma disappearance of infused Hb is relatively rapid, with a plasma half-disappearance time ranging from two to four hours depending on the experimental design, the preparation of Hb infused, and the animal species used (2,9,16). In our studies (9), after exchange transfusion of rats with crystalline Hb at levels of plasma Hb of 6-7 g/dl, we found, as mentioned before, a linear disappearance of free Hb with a plasma half-disappearance time of 3.5 hours. Although Hb solutions, such as that prepared by the crystallization method (20), have several potential applications as resuscitation fluids, the two limitations represented by the higher oxygen affinity and shorter intravascular retention time of free Hb as compared to Hb present in red cells could prevent a more generalized use of these solutions in fluid replacement therapy.

In our investigations we have found that the disadvantages of the short vascular retention time of Hb could be overcome by restoring plasma Hb and blood volume with repeated bolus injections or with continuous infusion of Hb solution after massive exchange transfusions with Hb solution (21). With this procedure the oxygen carrying capacity and the oncotic pressure of the vascular fluid are also maintained. This procedure could be adapted easily in an urban environment, but it would present serious difficulties in battlefield situations for the immediate attention to combat casualties. Therefore, to overcome the two Hb limitations mentioned above, we investigated the modification of the Hb molecule by subjecting this compound to pyridoxalation and subsequent polymerization reactions (31,32).

Pyridoxylation, done essentially by the method of Benesch et al. (33), yielded a product with a P_{50} ranging from 23 to 26 mm Hg. The pyridoxalated Hb was then polymerized with glutaraldehyde. Figure 5 presents a schematic illustration of the final procedure used for preparation of pyridoxalated-polymerized Hb, the appropriate concentrations of the various reagents and the P_{50} and n-value (Hill's coefficient) of intermediate and final products obtained in four preparations. The methemoglobin content of the final solutions of pyridoxalated-polymerized Hb was less than 5 g per 100 g of total Hb. Figure 6 illustrates the oxygen dissociation curves of unmodified Hb, pyridoxalated-polymerized Hb and whole blood. When compared to unmodified Hb, the curve of modified Hb is shifted to the right, indicating a lower oxygen affinity. Absorbancy curves between 350 and 650 nm of unmodified and of pyridoxalated-polymerized Hb show identical spectra. They represent the spectrophotometric curve of oxyhemoglobin, with absorption maxima at 414, 540 and 578 nm and absorption minima at 510 and 560 nm. Cellulose acetate electrophoresis resolution and column chromatographic analysis of preparations of Hb and of pyridoxalated-polymerized Hb indicated that 80 percent of Hb had been modified.

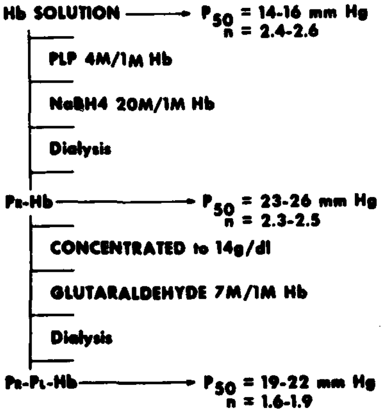

Hb SOLUTION
P_{50} = 14-16 mm Hg
n = 2.4-2.6
PLP 4M/1M Hb
NaBH4 20M/1M Hb
Dialysis
Pr-Hb
P_{50} = 23-26 mm Hg
n = 2.3-2.5
CONCENTRATED to 14g/dl
GLUTARALDEHYDE 7M/1M Hb
Dialysis
Pr-PL-Hb
P_{50} = 19-22 mm Hg
n = 1.6-1.9

Fig. 5. Schematized procedure for pyridoxalation and polymerization of crystalline hemoglobin. Hb= hemoglobin, PLP=pyridoxal phosphate, PR-Hb=pyridoxalated hemoglobin, PR=PL-HB=pyridoxalated-polymerized hemoglobin.

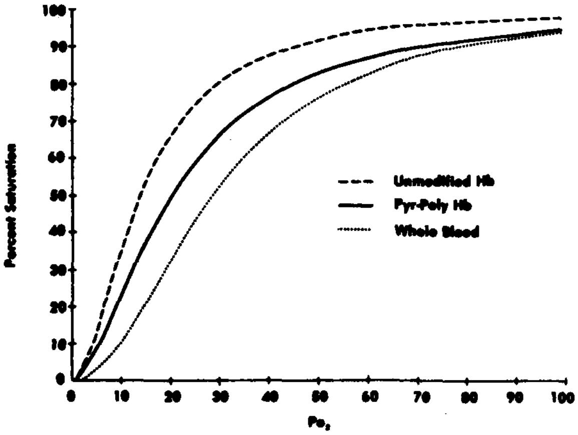


Fig. 6. Oxygen dissociation curves of unmodified Hb, pyridoxalated-polymerized Hb (Pyr-Poly Hb) and human blood.

The modified Hb was tested in vitro for coagulation activities. The results (Table 3) demonstrated that in the conventional coagulation tests used (TT, aPTT, PT, KCT) no pro- or anticoagulant activity was shown by the modified Hb when compared to a saline control.

TABLE 3

COAGULATION ACTIVITY*

	TT	aPTT	PT	KCT
Saline (control)	22.5 ± 0.1	29.8 ± 0.6	12.9 ± 0.1	83.8 ± 1.2
PR-PL-Hb	22.3 ± 0.3	30.2 ± 0.3	13.1 ± 0.3	85.8 ± 4.1

*TT = thrombin time, aPTT = activated partial thromboplastin time, PT = prothrombin time, KCT = kaolin coagulation time. Time in seconds. PR-PL-Hb = pyridoxalated-polymerized hemoglobin.

**Mean ± S.D. (n = 3)

The pyridoxalated-polymerized Hb was evaluated in vivo by exchange-transfusing rats to total blood replacement and the results were compared to those obtained using unmodified Hb. The characteristics of the Hb and pyridoxalated-polymerized Hb solutions used for exchange-transfusions are shown in Table 4. The two solutions are essentially the same in their composition except for the methemoglobin content which is slightly higher

TABLE 4
ANALYSIS OF HEMOGLOBIN (Hb) AND PYRIDOXALATED-POLYMERIZED HEMOGLOBIN (Pr-Pl-Hb) SOLUTIONS USED FOR EXCHANGED TRANSFUSIONS IN THE RAT

Assays	Hb Solution	Pr-Pl-Hb Solution
Hb (gm/dl)	7.1	7.4
MetHb (gm/100 gm Hb)	2.7	4.9
Na (mEq/L)	135	142
K (mEq/L)	0.4	0.3
Osmolality (mOs/Kg)	290	302
P_{50} (at pH 7.4)	14.3	20.6
n value	2.4	1.8
O_2 (volume percent)	9.8	10.2
Sterility	sterile	sterile

in the modified Hb solution. The two parameters, P_{50} and n-value, are different since they reflect characteristics specific for the two hemoglobins used. The results of total blood exchange in two groups of eight rats each, using unmodified Hb (control group) or pyridoxalated-polymerized Hb (experimental group) solution are shown in Figure 7.

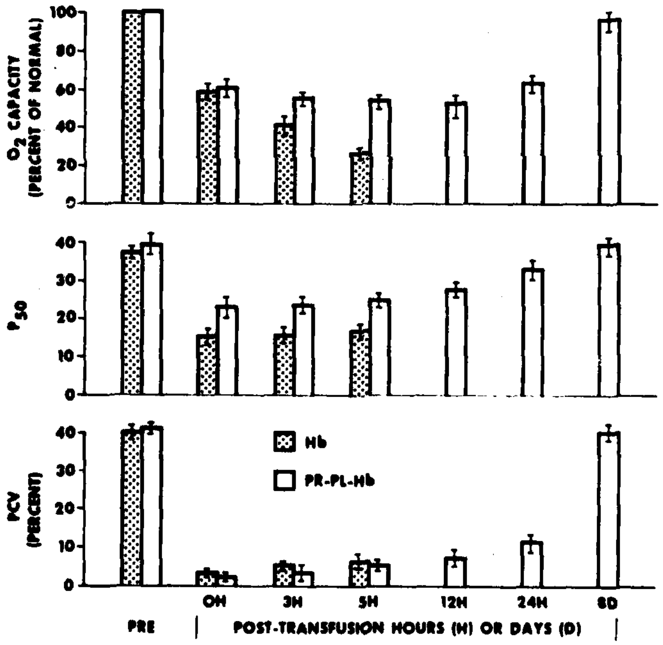


Fig. 7. Packed cell volume (PCV), P_{50} (at pH 7.4) and oxygen capacity before and after exchange-transfusions to 93-95 percent blood in the rat with hemoglobin (Hb) or pyridoxalated-polymerized hemoglobin (Pr-Pl-Hb) solution. Mean values ± SD (n=8).

All animals transfused with Hb solution died at approximately five hours after transfusion, whereas all animals transfused with pyridoxalated-polymerized Hb solution survived. Immediately after transfusion the hematocrit level reached 2-3 percent in the two groups decreasing from a pre-transfusion value of 40-41 percent and representing a 93-95 percent blood replacement. In the first 5 hours after transfusion both animal groups showed a slight increase in the hematocrit level, which in the modified Hb-treated rats continued to increase gradually afterwards, reaching a normal level at 8 days, due to regeneration of red cells. The P_{50} of the vascular fluid in the rats transfused with pyridoxalated-polymerized Hb solutions was signficantly greater than that observed in the animals transfused with unmodified Hb in the 5 hours after transfusion. In the experimental group the P_{50} increased after 5 hours and showed a normal value at 8 days after transfusion. Statistically significant differences were observed in the oxygen carrying capacity of the circulating fluid in the two groups of rats at 3 and 5 hours after transfusion. Although immediately after transfusion the oxygen carrying capacity was essentially the same in the two groups of animals, in the subsequent 5 hours it decreased to a value of 45 percent in the control group and only to 87 percent in the experimental group, from the level observed immediately after transfusion. In this latter group the amount of oxygen carried in the circulating fluid increased at 24 hours and showed a normal value at 8 days after transfusion. The reason for these significant differences in the oxygen carrying capacity in the two groups of animals is

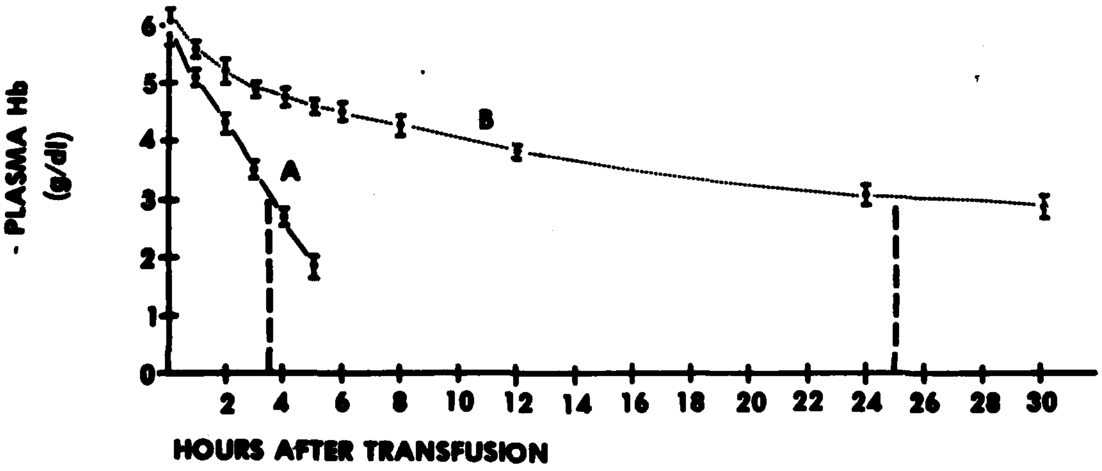


Fig. 8. Plasma disappearance of hemoglobin in rats exchanged with hemoglobin (A) or pyridoxalated-polymerized hemoglobin (B) solution to 93-95 percent blood replacement. The dash lines indicate time of 50 percent disappearance. Mean ± SD (n=4).

found in the rate of disappearance of plasma Hb which is shown in Figure 8 for two groups of four rats each, control and experimental group. The data in Figure 8 demonstrate that the plasma disappearance of Hb was faster in the Hb-treated than in the modified Hb-transfused animals. The plasma half-disappearance time, that is the time necessary for the plasma Hb concentration to reach one half the value observed at the end of exchange-transfusion was 3.5 hours for the control groups and 25 hours for the experimental group.

It appears, therefore, that the two limitations of higher oxygen affinity and lower intravascular retention of Hb can be overcome by the modification of the Hb molecule. The administration of a solution of pyridoxalated-polymerized Hb appears to be beneficial in total blood replacement since it promotes survival of the animals. We have investigated morphological effects on various organs, and preliminary studies have demonstrated that kidney and liver cell structures remain normal; however, other potential effects which may occur after massive transfusion with pyridoxalated-polymerized Hb solution must be studied in order to evaluate the significant potential value of this solution as a blood substitute.

CONCLUSIONS

The development and evaluation of an effective Hb solution as a blood substitute are important not only for the care of casualties resulting from mass disasters, but also for eventual use in other special clinical situations. Substantial improvements have been made by several recent research efforts in the general purity and quality of experimental Hb solutions. Such improvments have enabled investigators to study the potential applications of Hb solutions in far less ambiguous manner than was previously possible. These studies have produced a reasonably clear picture of the limitations of the current product and an insight for approaches in a systematic improvement.

The Hb solution as presently formulated, although has the limitation of short vascular retention time, potentially could be useful in several applications such as transfusions in patients who cannot receive immediate medical assistance, but could receive blood after a few hours, or transfusions in patients who cannot receive medical care for prolonged periods of time, provided that the Hb and blood volume losses are restored by periodic or continuous infusion of Hb solution. Furthermore, the present product could be useful in open heart surgery, in organ perfusion and in situations of uncontrolled bleeding.

It is hoped that research studies involved with modification of Hb, aimed at maintaining the tetrameric molecule in vivo, may provide a stable Hb compound having longer intrascular life as well as lower oxygen affinity and that the remaining issues could be clarified so that this blood substitute can be beneficial to man.

REFERENCES

1. Amberson WR: Blood substitutes. Biol Rev 12:48-86, 1937.
2. DeVenuto F, Friedman HI, Neville JR, Peck CC: Appraisal of hemoglobin solution as a blood substitute. Surg Gynecol Obstet 149:417-436, 1979.
3. Gruber VF: Blood replacement. Springer-VERLAC, Berlin, 1969.
4. Huestis DW, Bove JR, Busch S: Practical blood transfusion. 2nd ed., Little, Brown, Boston, 1976.
5. Bonhard K: Acute oxygen supply by infusion of hemoglobin solutions. Fed Proc 34:1466-1467, 1975.
6. Kaplan HR, Murthy VS: Hemoglobin solution: a potential oxygen transporting plasma volume expander. Fed Proc 34:1461-1465, 1975.
7. Moss, GS, Der Woskin R, Rosen AL: Transport of oxygen and carbon dioxide by hemoglobin-saline solution in the red cell-free primate. Surg Gynecol Obstet 142:357-362, 1976.
8. Peskin GW, O'Brien K, Rabiner SF: Stroma-free hemoglobin solution: The "ideal" blood substitute? Surgery 66:185-191, 1969.
9. DeVenuto F, Moores WY, Zegna AI, Zuck TF: Total and partial blood exchange in the rat with hemoglobin prepared by crystallization. Transfusion 17:555-562, 1977.
10. Amberson Wr, Flexner J, Steggerada FR, Mulder AG, Tendler MJ, Pancratz DS, Laug EP: On use of Ringer-Locke solutions containing hemoglobin as a substitute for normal blood in mammals. J Cell Comp Physiol 5:359-382, 1934.
11. DeVenuto F, Zegna AI: Plasma oncotic pressure during and after blood exchange with crystalline hemoglobin solution. Surg Gynecol Obstet 147:379-384, 1978.
12. Usami S, Chien S, Gregersen MI: Hemoglobin solution as a plasma expander: Effects on blood viscosity. Proc Soc Exp Biol Med 136:1232-1235, 1971.
13. DeVenuto F, Busse KR, Zegna AI: Viscosity of human blood hemodiluted with crystalline hemoglobin solution. Transfusion 21:752-756, 1981.
14. DeVenuto F: Stability of crystalline hemoglobin solution during extended storage. J Lab Clin Med 92:946-952, 1978.
15. Hamilton PB, Farr LE, Hiller A, Van Slyke DD: Preparation of hemoglobin solutions for intravenous infusion. J Exp Med 86:455-463, 1947.
16. Rabiner SF, Helbert JR, Lopas H, Friedman LH: Evaluation of stroma-free hemoglobin solution for use as a plasma expander. J Exp Med 126:1127-1142, 1967.

17. Cockin A, Das Gupta TK, Der Woskin R, Moss GS: Immunogenic properties of stroma versus stroma-free hemoglobin solutions. Surg Forum 23:19-21, 1972.
18. Sasazuki T: Studies on haptoglobin. I. Immunochemical properties of haptoglobin and antihemoglobin antibody. Proc Jap Acad 46:820, 1970.
19. DeVenuto F, Zegna AI, Busse KR: Lyophilization of crystalline hemoglobin solution and exchange transfusions with lyophilized, reconstituted hemoglobin. Surg Gynecol Obstet 148:69-75, 1979.
20. DeVenuto F, Zuck TF, Zegna AI, Moores WY: Characteristics of stroma-free hemoglobin prepared by crystallization. J Lab Clin Med 89:509-516, 1977.
21. DeVenuto F, Friedman HI, Mellick PW: Massive transfusions with crystalline hemoglobin solutions and subsequent replacement of hemoglobin and blood volume. Surg Gynecol Obstet 151:361-365, 1980.
22. DeVenuto F, Busse KR, Zegna AI: Oxygen transport by human blood hemodiluted with crystalline hemoglobin solution. Surg Gynecol Obstet 153:332-336, 1981.
23. Zuck TF, DeVenuto F, Neville JR, Friedman HI: Oncotic and oxygen transport effects of hemoglobin solutions. Prog Clin Biol Res 19:111-147, 1978.
24. Friedman HI, DeVenuto F, Zuck TF, Mellick P, Lollini L: Histologic and ultrastructural effects of stroma-free hemoglobin solution in rat liver, kidney, and brain. Surg Forum 28:3-5, 1977.
25. Friedman HI, DeVenuto F, Lollini L, Mellick P, Zuck TF: Morphological effects following massive exchange transfusions with a stroma-free hemoglobin solutions. I. Liver. Lab Invest 39:167-177, 1978.
26. Friedman HI, DeVenuto F, Lollini L, Mellick P, Zuck TF: Morphological effects following massive exchange transfusions with a stroma-free hemoglobin solution. II. Kidney. Lab Invest 40: 655-667, 1979.
27. Moores WY, DeVenuto F, Heydorn WH, Weiskopf RB, Baysinger M, Hannon JP: Improved porcine myocardial performance during severe anemia using a stroma-free hemoglobin solution. Fed Proc 39:2331, 1980.
28. Moores WY, DeVenuto F, Heydorn WH, Weiskopf RB, Baysinger BS, Greenburg AF, Utley JR: Extending the limits of hemodilution on cardiopulmonary bypass using stroma-free hemoglobin solution. J Thorac Cardiovasc Surg 81:155-162, 1981.
29. Nees JE, Hauser CJ, Shippy C, State D, Shoemaker WC: Comparison of cardiorespiratory effects of crystalline hemoglobin, whole blood, albumin and Ringer's lactate in the resuscitation of hemorrhagic shock in dogs. Surgery 83:639-647, 1978.
30. Messmer K, Jesch F, Schaff J, Schoenberg M, Peilsticker K, Bonhard K: Oxygen supply by stroma-free hemoglobin. Prog Clin Biol Res 19:175-190, 1978.

31. DeVenuto F, Zegna AI: Transfusions with pyridoxalated-polymerized hemoglobin solution. Transfusion 21:599, 1981.
32. DeVenuto F: Hemoglobin solutions as oxygen-delivering resuscitation fluids. Crit Care Med (in press).
33. Benesch RE, Benesch R, Renthal RD, Maeda N: Affinity labeling of the polyphosphate binding site of hemoglobin. Biochemistry 11:3576-3582, 1972.

FIELD PRODUCTION OF PURIFIED STERILE WATER FROM AVAILABLE WATER SOURCES BY USING A PORTABLE APPARATUS (U)

*FRANK DEVENUTO, PH.D.
ANGELO ZEGNA

LETTERMAN ARMY INSTITUTE OF RESEARCH
PRESIDIO OF SAN FRANCISCO, CA 94129

The development of portable equipment for the production of pure water from fresh pond, sea water, or other sources is important for military requirements in field situations. Also, it is necessary for many civilian activities whenever pure water is not available. The criteria for purity are influenced by the intended use of the water. Lake water which could be considered pure for swimming may not be suitable for drinking, and municipal water found pure for drinking is not considered pure for pharmaceutical preparations or for clinical use. The American Chemical Society, the American Society for Testing Materials, and the College of American Pathologists have specified various parameters for the purity of water. However, each of these agencies has proposed different standards (1), which apply for different purposes, and therefore, do not represent a uniform guide.

According to the criteria of the United States Pharmacopeia (USP) (2), sterile water for injection is a clear, colorless, odorless liquid; it is sterile without addition of antimicrobial agent or other substances, is pyrogen-free, and has a total solids content of 2 to 4 mg percent. In the usual urban environment, sterile water for injection is readily available; however, in field situations or in areas remote from supply sources, the procurement of purified water could present logistic difficulties. The problems associated with transportation, storage, and supply of large quantities of pure sterile water can be overcome by its production *in situ* by using any available source of water which is purified by a simple inexpensive process. Purified, injectable water may be needed in military field situations as a diluent for pharmaceutical formulations, for preparation of sterile saline solutions, for reconstitution of lyophized hemoglobin for fluid therapy (3-5), for solubilization of preservatives to be added to human blood for prolongation of its shelf life (6), for preparation of solutions for the rejuvenation of outdated red blood cells (7) or for other purposes.

The data presented here represent the results of the evaluation of a portable, compact, single unit apparatus used for the purification of sea water, pond water, and human urine, and the production of purified water that appears to satisfy the criteria for USP grade water for injection except for slightly higher content of total solids. Human urine was used in these studies to evaluate the potential of the purification process in extreme situations, such as a desert environment, where urine might be the only available source of water.

PURIFICATION OF WATER BY A PORTABLE APPARATUS

Three different sources of water were used for purification. Sea water was collected from the Marina at Fort Baker, California in the San Francisco Bay Area. Pond water was taken from the lagoon in the front of the Palace of Fine Arts in San Francisco. Pooled human urine was collected during a 24-hr period from several male laboratory workers. The apparatus used for the purification of water from the three different sources is manufactured by Allied Water Corporation, San Francisco, California. This portable SweetWater™ System, Model 200, enclosed in a fiberglass case, measures 45.7 x 78.7 cm, has a weight of 62 kg, and an output of 757 liters of water per day. The system is equipped with a water pump which, in the set-up used for these experiments, pumps the source water through three serial filters made of pure bleached cotton, cellulose, and activated carbon, then through a reverse osmosis purifier consisting of an acetate micropore filter, followed by a filter of activated carbon, a source of ultraviolet light, and finally, through an outlet provided with a sampling device. A

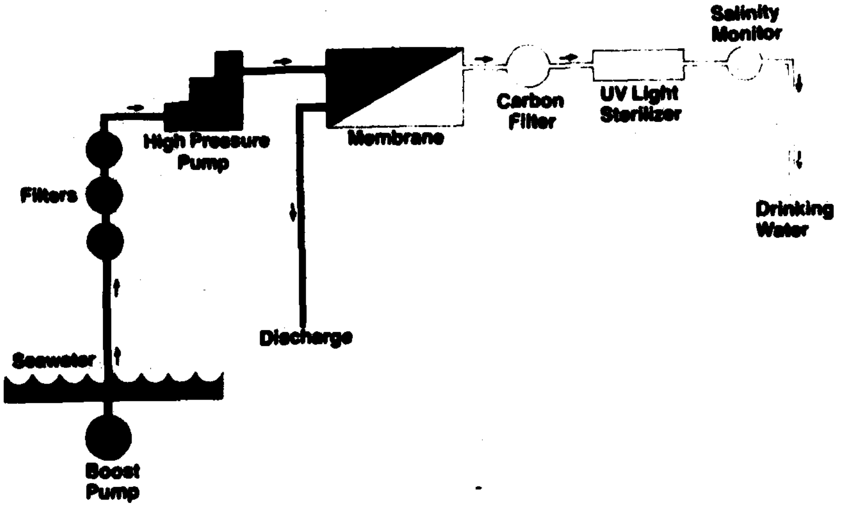


Fig. 1. Flow diagram of water purification process.

flow diagram of the water purification process is shown in Figure 1. In some experiments, a millipore Twin-90 sterile 0.22 μ filter was connected to the outlet system prior to collection of purified water to remove bacterial contamination.

The acetate micropore filter incorporated in the reverse osmosis purifier lasts for two years or longer if properly maintained. It is continuously rinsed by water pumped into the apparatus, thus preventing the accumulation of particulate matter on the filter. The three serial filters placed before the reverse osmosis purifier and the activated carbon filter which follows remain effective for a period of time dependent on the impurities present in the water source used. Purification of sea water in a continuous operation requires replacement or regeneration after one week for the first-in-line filter and after two weeks for the other filters in order to obtain purified water of consistent high quality. The apparatus is portable, and can be used anywhere water is needed as long as a water supply of some source is available. It can be modified to use different power sources such as a combustion engine or even manpower.

RESULTS

The ion content of the water samples from sea water, fresh pond water, and human urine before and after purification are shown in Table I. The

TABLE I
ION CONTENT* BEFORE AND AFTER WATER PURIFICATION

	USP	SEA		POND		URINE	
		Before	After	Before	After	Before	After
Na (mg/dl)†	0.0	1,023.05	18.39	19.54	1.15	188.52	2.30
K (mg/dl)†	0.0	39.10	0.78	2.35	0.39	174.00	1.17
Ca (mg/dl)†	0.2	22.80	0.31	2.65	0.05	8.30	0.30
Mg (mg/dl)†	0.59	14.70	0.28	1.12	0.00	4.52	0.35
Cl (mg/dl)†	0.0	1,730.11	13.26	9.08	0.00	5.67	1.06
P (mg/dl)†	0.0	0.0	0.0	0.95	0.20	105.20	0.30

*Values of 0.0 represent levels below the limits of detection, which in mg/dl are: 1.1 for Na, 0.08 for K, 0.9 for Cl and 0.05 for Ca, Mg, and P.

†ppm = data in Table x 10

analysis of a sample of USP water is also included in this Table for comparison. Sea water with a high Na and Cl content, as expected, is purified to the extent that, with a single passage through the purification system, more than 98% of Na and more than 99% of Cl are removed. Na, K, and P, which are present in elevated amounts in human urine, are also reduced by the purification procedure to 1 to 2% of the initial value. All other ions indicated in the Table are reduced considerably independently of the source of water used. The detection limits of the methods used are indicated in Table I.

Table II shows the metal content in the water samples before and after purification. In some samples the metal content before purification was below the limit of sensitivity of the assay (Table II). However, in those instances where significant amounts were present, removal was achieved by the purification procedure, except for zinc in the urine sample where a decrease of 87.5% was observed.

TABLE II
METAL CONTENT BEFORE AND AFTER WATER PURIFICATION

	USP	SEA		POND		URINE	
		Before	After	Before	After	Before	After
Cu (ppm)	<0.06	<0.06	<0.06	<0.06	<0.06	0.71	<0.06
Fe (ppm)	<0.11	0.38	<0.11	<0.11	<0.11	0.34	<0.11
Mn (ppm)	<0.06	0.07	<0.06	<0.06	<0.06	<0.06	<0.06
Zn (ppm)	0.024	0.059	0.02	0.043	<0.011	0.526	0.066

Conductivity, electrical resistance, osmolality, and pH of the water samples before and after purification are shown in Table III. With a considerable decrease of ion content, as observed in Table I, a corresponding decrease in electrical conductance and parallel increase in electrical resistance was obtained, as expected. The decrease in osmolality after purification also reflects the removal of osmotic material from the water of different sources. The difference observed in the pH of pond water before and after purification may indicate loss of CO_2 dissolved in fresh pond water and/or removal of other alkaline material.

In Table IV, the spectrophotometric absorbance between 220 and 650 nm is indicative of the presence of organic or pigmented material with light

TABLE III
CONDUCTIVITY, ELECTRICAL RESISTANCE, OSMOLALITY, AND pH
BEFORE AND AFTER WATER PURIFICATION

	USP	SEA		POND		URINE	
		Before	After	Before	After	Before	After
Conductivity (μmhos)	2.8	37,000	750	1,000	29	11,000	200
Electrical Resistance (ohms)	400,000	30	1,400	1,000	37,000	95	5,200
Osmolality (mOsm/kg)	1	925	15	18	3	426	33
pH	7.45	7.20	7.60	9.2	6.30	6.45	6.20

TABLE IV
ABSORBANCE, FLUORESCENCE, TOTAL PROTEIN, AND TOTAL
MATTER BEFORE AND AFTER WATER PURIFICATION

	USP	SEA		POND		URINE	
		Before	After	Before	After	Before	After
Absorbance (220-650nm)	0.0	0.0	0.0	220-280nm	0.0	280nm	0.0
Fluorescence (Iu 340/465*)	0.0	16.5	0.0	>100	0.0	>100	0.0
Total Protein (g/dl)	---	---	---	---	---	0.15	0.0
Total Matter (g/dl)	0.003	4.05	0.034	0.17	0.007	3.02	0.20

*Intensity units at 340nm excitation and 465nm emission.

absorbing characteristics in the ultraviolet or visible region. Pond water showed absorbance in the 220 to 280 nm region and uring at 280 nm before purification. After purification, the light absorbing material was removed since no absorbance was observed. Significant fluorescence was measured in sea water, and to a greater extent, in pond water and urine before purification. However, this fluorescence was not present in samples of purified water. No significant amounts of protein could be detected in the sea or pond water; the protein content determined in the urine sample was absent after the process of purification. The residue remaining after evaporation of 50 ml of water from the three different sources before and after purification is shown quantitatively in Table IV and is illustrated in Figure 2 (for sea water), Figure 3 (for pond water), and Figure 4 (for human urine).

Figure 5 depicts the results of the gas chromatographic analysis of samples of sea water and urine before and after purification and of a

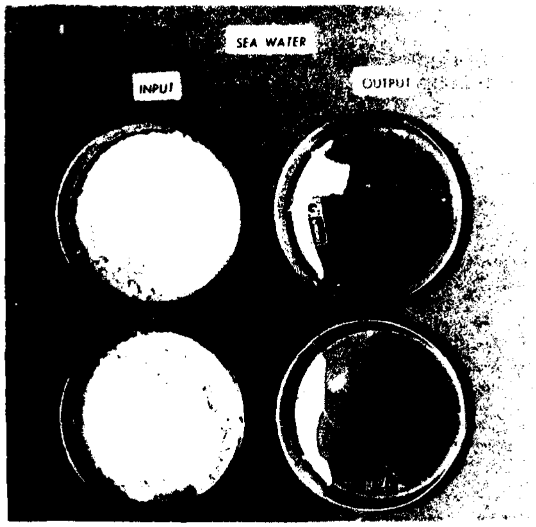

SEA WATER
INPUT
OUTPUT

Fig. 2. Residue after evaporation of water from 50 ml sea water before and after purification. In duplicate.

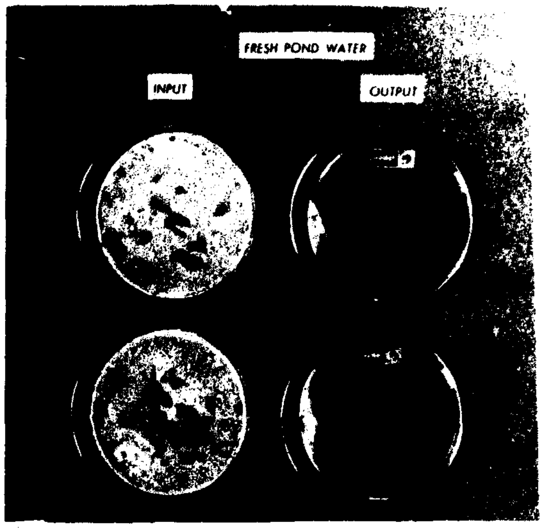

FRESH POND WATER
INPUT
OUTPUT

Fig. 3. Residue after evaporation of water from 50 ml fresh pond water before and after purification. In duplicate.

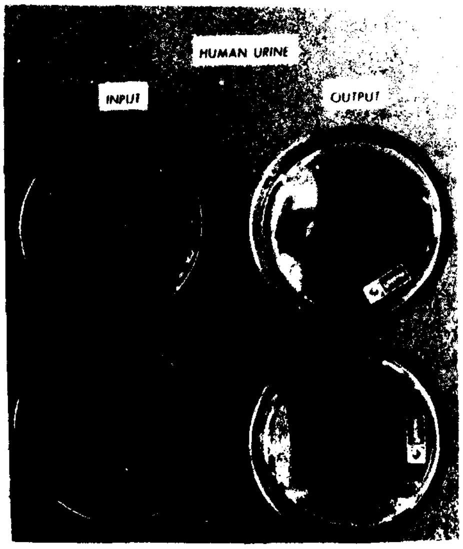

HUMAN URINE
INPUT
OUTPUT

Fig. 4. Residue after evaporation of water from 50 ml human urine before and after purification. In duplicate.

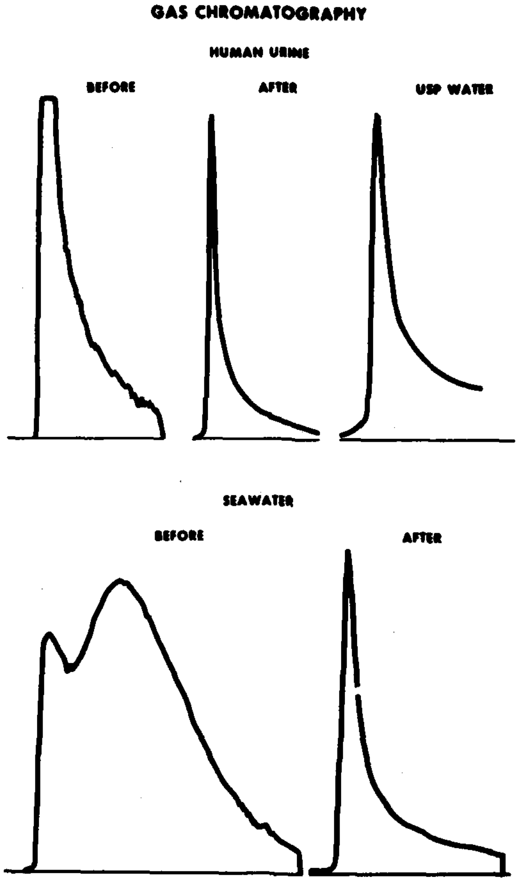


Fig. 5. Gas chromatographic analysis of sea water and human urine before and after purification. A pattern of sample of USP water is also illustrated. The ordinate indicates the frequency signals of the chromatographic detector and the abcissa the retention time of the gas products.

sample of USP water. Sea water or urine, prior to purification, shows patterns indicative of the presence of several impurities, but after purification the pattern is similar to that obtained with USP water. Purified water obtained from the three different sources appears as a clear, colorless, odorless liquid.

All the data obtained in Tables I to IV and Figures 2 to 5 are essentially the same with or without addition of the Twin-90, 0.22 μ millipore filter prior to collection of the purified samples. The addition of this sterile filter, however, has a remarkable effect on the bacteriological

TABLE V
BACTERIAL ANALYSIS. COLONY FORMING UNITS PER ML.

	Before Purification	After Purification	
		Without Twin-90	With Twin-90
Sea Water	1.1*	1.2**	0.0
Pond Water	4.5 x 10^3***	3.1 x 10^3***	0.0
Urine	3.0 x 10^4†	1.2‡	0.0

Bacteria Identified:

**Klebsiella pneumoniae*, *Enterobacter agglomerans*, *Serratia liquefaciens*

***Enterobacter agglomerans* and *Serratia liquefaciens*

***Proteus

†Bacillus, *Escherichia coli*, *Pseudomonas aeruginosa*, *Pseudomonas fluorescans*, *Proteus mirabalis*, *Enterobacter agglomerans*

‡*Pseudomonas*

TABLE VI
DETECTION OF PYROGENIC BACTERIAL ENDOTOXIN BEFORE AND AFTER WATER PURIFICATION

		Reaction Duplicate		Result
positive control	50.0 ng/ml	+	+	positive
(*E. coli* endotoxin)	0.5	+	+	positive
	0.1	+	+	positive
	0.05	±	±	positive
	0.025	-	-	negative
	0.0125	-	-	negative
	0.006	-	-	negative
Control (USP water)		-	-	negative
Sea water (before)		±	±	positive
(after, without Twin-90)		±	±	positive
(after, with Twin-90)		-	-	negative
Pond water (before)		+	+	positive
(after, without Twin-90)		±	±	positive
(after, with Twin-90)		-	-	negative
Urine (before)		+	+	positive
(after, without Twin-90)		±	±	positive
(after, with Twin-90)		-	-	negative

assays. The results of these tests are shown in Table V. Water obtained from the three different sources is contaminated by different bacteria with very heavy contamination observed in pond water and urine. Purification without addition of the sterile filter shows little or no effect on the elimination of bacterial contamination from sea or pond water, although bacteria present in human urine are considerably reduced. However, the addition of the sterile filter prior to collection of water assures absence of bacterial contamination in the purified water.

Table VI shows the results of the detection of pyrogenic bacterial endotoxin in samples of water before and after purification. The limit of sensitivity of the test is indicated by data of the positive control, represented by Escherichia coli endotoxin; this limit is reached at a concentration of 0.05 ng/ml. Based on earlier assays, a pyrogenic response was obtained in rabbits at a concentration of 0.5 ng/ml. Prior to purification, water from the three different sources gave a positive reaction which is present also in purified water samples obtained without addition of a sterile filter. However, all the water samples collected after filtration through the sterile filter show a negative reaction, indicating removal of pyrogenic bacterial endotoxin. A negative reaction was observed also in the USP control water sample.

CONCLUSIONS

The data obtained in our studies demonstrate that the purification system used satisfies all criteria for USP grade injectable water, except one: the limits of total solids. However, the purified water is nonpyrogenic, clear, colorless, odorless, and it is sterilized without addition of antimicrobial agents or other substances. The ion, metal content, and organic material present in human urine, pond water, and sea water are reduced considerably or completely removed by a single passage through the system. Although the total solids in the purified water exceed the limit of 2 to 4 mg% established for USP water, sodium and chlorine ions represent 96% and 64% of the total solids present in water purified from sea water and urine, respectively. However, these ions are generally added (900 mg sodium chloride per 100 ml) in order to make isotonic water for injection. The results indicate also that organic substances, with spectrophotometric absorption in the ultraviolet and visible regions, fluorescent compounds and protein material present in the water source are removed by the purification process. It is important to emphasize that this purification system, as presently available, will not manufacture water completely sterile and free of pyrogenic bacterial endotoxins, but requires the addition of a small porosity, sterile, in-line filter to produce water without bacterial contamination.

It appears that the reverse osmosis process is efficient in the production of purified water and it has been utilized for the preparation of drinking water in a desalinization project (8). The system evaluated in these studies represents a multipurpose water purification process which could be useful in many non-military applications such as disaster sites where sources of water are contaminated, at construction sites, off-shore drilling platforms, on commercial and pleasure boats, and other activities where pure water is needed but not available.

REFERENCES

1. Dharan M: Standards and methods to obtain reagent grade water. Lab Manag 16:39-42, 1978.
2. The United States Pharmacopeia, USP Convention, Inc., Rockville, MD, 19th revision, p. 540, 1975.
3. DeVenuto F, Moores WY, Zegna AI, Busse KR: Total and partial blood exchange in the rat with hemoglobin prepared by crystallization. Transfusion 17:555-562, 1977.
4. DeVenuto F, Zegna AI, Busse KR: Lyophilization of crystalline hemoglobin solution and exchange transfusions with lyophilized-reconstituted hemoglobin. Surg Gynecol Obstet 148:69-75, 1979.
5. DeVenuto F, Zuck TF, Zegna AI, Moores WY: Characteristics of stroma-free hemoglobin prepared by crystallization. J Lab Clin Med 89:509-516, 1977.
6. Moore GL, Ledford ME, Brummell MR, Brooks DE: The potential use of dihydroxyacetone for improved 2,3-DPG maintenance in red cell storage. Transfusion 20:24-31, 1980.
7. Valeri CR, Zaroulis CG: Rejuvenation and freezing of outdated stored human red cells. N Engl J Med 287:1307-1313, 1972.
8. Schmitt, RP: Military requirements for water supply. First Desalination Congress of the American Continent 1:1-8 (V-2), 1976.

A NEW ANALYSIS OF THE UNWINDING RIBBON AS A DELAYED ARMING DEVICE (U)

*WILLIAM P. DUNN, Ph.D.
US ARMY LARGE CALIBER
WEAPON SYSTEMS LABORATORY
DOVER, NJ 07801

1. INTRODUCTION

The objective of this work was to formulate theoretical methods to enable engineers to design unwinding ribbons for use as delay arming mechanisms with reasonable accuracy and a minimum development effort. The unwinding ribbon considered here is a "wrapped" spring, which is a spiral spring made from flat metal stock closely wound. In the unstressed condition all the coils of the spring are touching. The results of the analysis are given and compared with the experimental results obtained by T.B. Alfriend*.

This is a more complete study than that of Alfriend since no assumptions are made concerning the moment of inertia of the coil and hub or the tension force in the ribbon bridge. Hence, two empirical constants in Alfriend's analysis were dropped in favor of exact expressions.

2. DESCRIPTION OF UNWINDER DEVICE

The basic components of the Unwinder device are shown schematically in Figure 1. The spring, A, is wrapped around and fastened at its inner end to the shaft, B. The outer end of the spring is fastened to the outer case, C, at the point D. The outer case, C, is fixed to and rotates with the projectile. The axis of the spring and of the inner shaft, B, are coincident with the longitudinal axis of the projectile as shown in Figure 1. Upon firing, torsional acceleration causes the spring to wind up tightly. After the torsional acceleration ceases, the centrifugal forces acting on the spring will tend to unwind it. During this unwinding process, the inner shaft, B, will rotate relative to the housing, and this motion can be used to close a switch, to rotate a firing pin in line with a detonator or to cause other arming processes.

---

*"Study of Wrapped Springs for Application to a Delayed Arming Device"- T.B. Alfriend - Summary Report ER-1404 Aircraft Armaments Inc.-1958.

3. ALFRIEND'S ANALYSIS

To write the differential equation of motion of the spring unwinding due to the action of centrifugal forces, Alfriend* assumed that the unbalanced torque on the inner shaft is a function of $K\omega^2$, where ω is the net angular velocity of the unbalanced spring mass and K is some experimentally determined constant. Pictorially, Alfriend represented the physical problem as shown in Figure 2.

The internal force vectors $\underset{\sim}{F}$ and $\underset{\sim}{M}$ are such that

$\underset{\sim}{M}$ = the torque required to deflect the spring statically from its initial unstressed state with radius r, to the radius of the outer case, r_{so} ;

$\underset{\sim}{F}$ = the tensile force in the spring due to centrifugal forces acting on the unbalanced length of the spring between points A and B, which are the points of contact between the stretched and coiled spring material at the inner and outer coils, respectively. The initial positions of these points are designated A_0 and B_0 .

The appropriate Euler differential equation of motion for the inner shaft and attached spring is then

$$Fr + M = I\ddot{\theta} \tag{1}$$

where θ is the relative angular displacement of the shaft with respect to the outer case. The moment of inertia, I, of the total revolving mass on the inner shaft is a function of the radius, r , which decreases as the spring unwinds. Alfriend assumed

$$I = I_0 + C\theta \tag{2}$$

where,

I_0 = the initial moment of inertia of the spring plus the inner shaft

C = assumed constant rate of decrease in I as the spring unwinds.

Alfriend then assumed

$$F = K\omega^2 \tag{3}$$

*See reference in Introduction.

where as aforementioned, K is experimentally determined and ω is the total angular velocity of the unbalanced spring length. He further assumed that

$$\omega = \omega_0 - \frac{r}{r_{20}} \dot{\theta} \tag{4}$$

where,

$$r = r_0 - \frac{\delta}{2\pi} \theta \tag{5}$$

and,

ω_0 = constant angular velocity of the outer case

δ = spring thickness

r_0 = initial outside radius of the spring on the inner shaft.

Substituting equations (2) - (5) into (1), the equation of motion becomes

$$K\left[\omega_0 - \frac{(r_0 - \frac{\delta}{2\pi}\theta)}{r_{20}} \dot{\theta}\right]^2 \left(r_0 - \frac{\delta}{2\pi}\theta\right) + M = (I_0 - c\theta)\ddot{\theta} \tag{6}$$

where,

$$M = \frac{1}{24} E b \delta^3 r_{10} \left(\frac{1}{r_{10}} - \frac{1}{r_{20}}\right)^2 \tag{6-A}$$

and,

E = the elastic modulus (Young's) of the spring material

b = the width of the spring material

r_{10} = the radius of the inner shaft

r_{20} = the radius of the outer case

4. MODIFIED ALFRIEND THEORY

A more accurate characterization of the behavior of the spring can be obtained by eliminating the assumptions made by Alfriend. To accomplish this, equations (2) and (3) are replaced by their analytically derived forms. This eliminates the two constants K and c introduced through the assumptions; in place of equation (2),

$$I = \frac{1}{2}\pi b\left[(\rho_{sw} - \rho_{sp}) r_w^4 + \rho_{sp} r^4\right] \tag{7}$$

where ρ_{sw} , ρ_{sp} are the inner shaft and spring densities, respectively.

To revise equation (3), consider Newton's 2nd Law of Motion i.e., the force is proportional to the change in momentum with time. and let

$m_{AB} = \rho_{sp}\delta b\sqrt{r_2^2 - r^2}$ = mass of the spring material between points "A" and "B" in Figure 3 (8)

$\underset{\sim}{r}_{cm} = \frac{1}{2}(\underset{\sim}{r}_2 + \underset{\sim}{r})$ = the position vector to the center of mass of material between the points "A" and "B" in Figure 3 (9)

$\underset{\sim}{v}_{cm} = \frac{d}{dt}(\underset{\sim}{r}_{cm})$ = the velocity vector of the mass center of the material between the points "A" and "B" in Figure 3 (10)

where

$\underset{\sim}{r}$ = the position vector to the last point of contact between the inner coiled spring material and the uncoiled spring material = $(r_0 - \delta/2\pi \cdot \theta)(-\sin\theta\,\hat{i} + \cos\theta\,\hat{j})$ (11)

$\underset{\sim}{r}_2$ = the position vector to the last point of contact between the outer coiled spring material and the uncoiled spring material $\underset{\sim}{r}_2 = (r_{20} - \delta/2\pi \cdot \alpha)[\cos(\lambda+\theta)\,\hat{i} + \sin(\lambda+\theta)\,\hat{j}]$ (12)

and

r_0 = the initial radius of the last point of contact between the inner coiled spring and the uncoiled spring.

r_{20} = the inner radius to the outer case

θ = the angle delineating the material unwrapped from the inner core

α = the angle delineating the material wrapped onto the outer case

λ = angle between the position vector $\underset{\sim}{r}_2$ and the uncoiled material = $\sin^{-1}(r/r_2)$

r = the scalar value of $|\underset{\sim}{r}| = r_0 - (\delta/2\pi)\theta$ (13)

$r_2 =$ the scalar value of $|\underset{\sim}{r_2}| = r_{20} - (\delta/2\pi)\alpha$ (14)

Thus, the force acting on the ribbon segment ($\overline{AB}$) is,

$$\underset{\sim}{F}_{AB} = \frac{d}{dt}(m_{AB}\underset{\sim}{V}_{cm}) = \frac{dm_{AB}}{dt}\underset{\sim}{V}_{cm} + m_{AB}\frac{d\underset{\sim}{V}_{cm}}{dt} \tag{15}$$

where,

$$\frac{dm_{AB}}{dt} = -\frac{\rho_{sp}\delta^2 b\,(r_2\dot{\alpha} - r\dot{\theta})}{2\pi\sqrt{r_2^2 - r^2}} \tag{16-A}$$

and in place of equation (6-A), we define

$$M = \frac{1}{24}Eb\delta^3 r\left(\frac{1}{r} - \frac{1}{r_2}\right)^2 \tag{16-B}$$

The transformation expressions relating the coordinate system fixed in the inner shaft to the ground reference coordinate system are:

$$\left.\begin{aligned}\hat{i} &= \cos\omega_b t\,\hat{I} - \sin\omega_b t\,\hat{J}\\ \hat{j} &= \sin\omega_b t\,\hat{I} + \cos\omega_b t\,\hat{J}\end{aligned}\right\} \tag{17}$$

$$\left.\begin{aligned}\hat{I} &= \cos\omega_b t\,\hat{i} + \sin\omega_b t\,\hat{j}\\ \hat{J} &= -\sin\omega_b t\,\hat{i} + \cos\omega_b t\,\hat{j}\end{aligned}\right\} \tag{18}$$

Substituting equations (17) into equations (9), (10) and $d\underset{\sim}{V}_{cm}/dt$ yields, upon simplification,

$$\underset{\sim}{r}_{cm} = \frac{1}{2}\left\{\left[r_2\cos(\lambda+\theta-\omega_b t) - r\sin(\theta-\omega_b t)\right]\hat{I} + \left[r_2\sin(\lambda+\theta-\omega_b t) + r\cos(\theta-\omega_b t)\right]\hat{J}\right\} \tag{19}$$

$$\begin{aligned}\underset{\sim}{V}_{cm} = \frac{1}{2}\Big\{&\left[\dot{r}_2\cos(\lambda+\theta-\omega_b t) - \dot{r}\sin(\theta-\omega_b t) - r_2(\dot{\lambda}+\dot{\theta}-\omega_b)\sin(\lambda+\theta-\omega_b t) - r(\dot{\theta}-\omega_b)\cos(\theta-\omega_b t)\right]\hat{I}\\ &+\left[\dot{r}_2\sin(\lambda+\theta-\omega_b t) + \dot{r}\cos(\theta-\omega_b t) + r_2(\dot{\lambda}+\dot{\theta}-\omega_b)\cos(\lambda+\theta-\omega_b t) - r(\dot{\theta}-\omega_b)\sin(\theta-\omega_b t)\right]\hat{J}\Big\}\end{aligned} \tag{20}$$

$$\begin{aligned}\frac{d\underset{\sim}{V}_{cm}}{dt} = \frac{1}{2}\Big\{&\left[\ddot{r}_2 - r_2(\dot{\lambda}+\dot{\theta}-\omega_b)^2\right]\cos(\lambda+\theta-\omega_b t) - \left[\ddot{r} - r(\dot{\theta}-\omega_b)^2\right]\sin(\theta-\omega_b t)\\ &-\left[2\dot{r}_2(\dot{\lambda}+\dot{\theta}-\omega_b) + r_2(\ddot{\lambda}+\ddot{\theta})\right]\sin(\lambda+\theta-\omega_b t) - \left[2\dot{r}(\dot{\theta}-\omega_b) + r\ddot{\theta}\right]\cos(\theta-\omega_b t)\Big\}\hat{I}\\ +\frac{1}{2}\Big\{&\left[\ddot{r}_2 - r_2(\dot{\lambda}+\dot{\theta}-\omega_b)^2\right]\sin(\lambda+\theta-\omega_b t) + \left[\ddot{r} - r(\dot{\theta}-\omega_b)^2\right]\cos(\theta-\omega_b t)\\ &+\left[2\dot{r}_2(\dot{\lambda}+\dot{\theta}-\omega_b) + r_2(\ddot{\lambda}+\ddot{\theta})\right]\cos(\lambda+\theta-\omega_b t) - \left[2\dot{r}(\dot{\theta}-\omega_b) + r\ddot{\theta}\right]\sin(\theta-\omega_b t)\Big\}\hat{J}\end{aligned} \tag{21}$$

Defining,

$$r_c \equiv \sqrt{r_2^2 - r^2} \qquad (22)$$

and noting from Figure (3) that

$$\lambda_0 + \alpha = \lambda + \theta \quad , \qquad (23)$$

differentiation then provides the relations

$$\dot{\alpha} = \dot{\lambda} + \dot{\theta} \quad ; \quad \ddot{\alpha} = \ddot{\lambda} + \ddot{\theta} \qquad (24)$$

Now, using equations (18) or (20) and (21); along with equations (22), (23) and (24), obtain from equation (15)

$$\begin{aligned}
F_{AB} = \tfrac{1}{2} \rho_w \delta b \Big[& \Big\{ \big[\ddot{r}_2 - r_2(\dot{\alpha} - \omega_b)^2\big] r_c - \frac{\delta \dot{r}_2 (r_2\dot{\alpha} - r\dot{\theta})}{2\pi r_c} \Big\} \cos(\lambda_0 + \alpha) \\
& - \Big\{ \big[2\dot{r}_2(\dot{\alpha} - \omega_b) + r_2\ddot{\alpha}\big] r_c - \frac{\delta r_2 (r_2\dot{\alpha} - r\dot{\theta})(\dot{\alpha} - \omega_b)}{2\pi r_c} \Big\} \sin(\lambda_0 + \alpha) \\
& - \Big\{ \big[2\dot{r}(\dot{\theta} - \omega_b) + r\ddot{\theta}\big] r_c - \frac{\delta r(\dot{\theta} - \omega_b)(r_2\dot{\alpha} - r\dot{\theta})}{2\pi r_c} \Big\} \cos\theta \\
& - \Big\{ \big[\ddot{r} - r(\dot{\theta} - \omega_b)^2\big] r_c - \frac{\delta \dot{r} (r_2\dot{\alpha} - r\dot{\theta})}{2\pi r_c} \Big\} \sin\theta \Big] \hat{i} \\
+ \tfrac{1}{2} \rho_w \delta b \Big[& \Big\{ \big[\ddot{r}_2 - r_2(\dot{\alpha} - \omega_b)^2\big] r_c - \frac{\delta \dot{r}_2 (r_2\dot{\alpha} - r\dot{\theta})}{2\pi r_c} \Big\} \sin(\lambda_0 + \alpha) \\
& + \Big\{ \big[2\dot{r}_2(\dot{\alpha} - \omega_b) + r_2\ddot{\alpha}\big] r_c - \frac{\delta r_2 (r_2\dot{\alpha} - r\dot{\theta})(\dot{\alpha} - \omega_b)}{2\pi r_c} \Big\} \cos(\lambda_0 + \alpha) \\
& - \Big\{ \big[2\dot{r}(\dot{\theta} - \omega_b) + r\ddot{\theta}\big] r_c - \frac{\delta r(\dot{\theta} - \omega_b)(r_2\dot{\alpha} - r\dot{\theta})}{2\pi r_c} \Big\} \sin\theta \\
& + \Big\{ \big[\ddot{r} - r(\dot{\theta} - \omega_b)^2\big] r_c - \frac{\delta \dot{r} (r_2\dot{\alpha} - r\dot{\theta})}{2\pi r_c} \Big\} \cos\theta \Big] \hat{j} \equiv F_1 \hat{i} + F_2 \hat{j}
\end{aligned} \qquad (25)$$

where F_1 and F_2 represent the scalar variables multiplying the unit vectors $\hat{i}$ and $\hat{j}$ respectively. From Figure 3, the tension in the segment $\overline{AB}$ is then

$$F = F_1 \cos\theta + F_2 \sin\theta \qquad (26)$$

Substituting F_1 and F_2 from equation (25) into (26) and making use of the trigonometric identities

$$\begin{aligned}
\sin(x \pm y) &= \sin x \cos y \pm \cos x \sin y \\
\cos(x \pm y) &= \cos x \cos y \mp \sin x \sin y
\end{aligned} \qquad (27)$$

yields, upon simplification,

$$F = \frac{1}{2}\rho_{sp}\delta b\Big[\Big\{[\ddot{r}_2 - r_2(\dot{\alpha}-\omega_0)^2]r_c - \frac{\delta \dot{r}_2(r_2\dot{\alpha} - r\dot{\theta})}{2\pi r_c}\Big\}\cos\lambda$$
$$+\Big\{[2\dot{r}_2(\dot{\alpha}-\omega_0)+r_2\ddot{\alpha}]r_c - \frac{\delta r_2(r_2\dot{\alpha}-r\dot{\theta})(\dot{\alpha}-\omega_0)}{2\pi r_c}\Big\}\sin\lambda$$
$$-\Big\{[2\dot{r}(\dot{\theta}-\omega_0)+r\ddot{\theta}]r_c - \frac{\delta r(\dot{\theta}-\omega_0)(r_2\dot{\alpha}-r\dot{\theta})}{2\pi r_c}\Big\}\Big] \tag{28}$$

After substituting (See Figure 3),

$$\sin\lambda = \frac{r}{r_2} \qquad , \qquad \cos\lambda = \frac{r_c}{r_2} \tag{29}$$

into equation (28) and then using this result in equation (1), the following expression is obtained.

$$\ddot{\theta} - \frac{M}{I} = A_0\Big[r_c\Big\{[\ddot{r}_2 - r(\dot{\alpha}-\omega_0)^2]r_c - \frac{\delta\dot{r}_2}{2\pi r_c}(r_2\dot{\alpha}-r\dot{\theta})\Big\} + r\Big\{[2\dot{r}_2(\dot{\alpha}-\omega_0)+r_2\ddot{\alpha}]r_c$$
$$-\frac{\delta r_2}{2\pi r_c}(r_2\dot{\alpha}-r\dot{\theta})(\dot{\alpha}-\omega_0)\Big\} - r_2\Big\{[2\dot{r}(\dot{\theta}-\omega_0)+r\ddot{\theta}]r_c - \frac{\delta r}{2\pi r_c}(\dot{\theta}-\omega_0)(r_2\dot{\alpha}-r\dot{\theta})\Big\}\Big] \tag{30}$$

where,

$$A_0 = \frac{\rho_{sp}\delta b r}{2 r_2 I} = \frac{\delta}{\pi r_2 r^3} \tag{31}$$

for the case $\rho_{sN} = \rho_{sp}$.

5. RELATIONSHIP BETWEEN WRAPPING ANGLE (α) AND UNWRAPPING ANGLE (θ):

To obtain α as a function of θ , note from Figure 3 that the material unwrapped from the inner shaft must equal the material wrapped on the outer case plus the material going into the increased length of the "bridge", $\overline{AB}$. Mathematically stated,

$$\int_0^{\theta} r\,d\theta' = \int_0^{\alpha} r_2\,d\alpha' + \left(\sqrt{r_2^2 - r^2} - \sqrt{r_{20}^2 - r_0^2}\right) \tag{32}$$

Substituting equations (13) and (14) into equation (32) and integrating yields,

$$\left(r_0 - \frac{\delta}{4\pi}\theta\right)\theta = \left(r_{20} - \frac{\delta}{4\pi}\alpha\right)\alpha + \left[\sqrt{\left(r_{20} - \frac{\delta}{2\pi}\alpha\right)^2 - \left(r_0 - \frac{\delta}{2\pi}\theta\right)^2} - \sqrt{r_{20}^2 - r_0^2}\right] \tag{33}$$

Expanding and simplifying equation (33) results in

$$\alpha^4-\frac{8\pi r_{20}}{\delta}\alpha^3+\frac{4}{\delta^2}\left[4\pi^2r_{20}^2+\delta^2-2\pi\delta\sqrt{r_{20}^2-r_{10}^2}-2\pi\delta\left(r_0-\frac{\delta}{4\pi}\theta\right)\theta\right]\alpha^2-\frac{16\pi^2r_{20}}{\delta^2}\left[\frac{\delta}{4\pi}\right.$$
$$\left.-2\sqrt{r_{20}^2-r_{10}^2}-2\left(r_0-\frac{\delta\theta}{4\pi}\right)\theta\right]\alpha+\frac{4\pi^2}{\delta^2}\left[\frac{4\pi\delta\theta}{\pi}-4\left(r_0-\frac{\delta\theta}{4\pi}\right)\theta^2\pm4\left(r_0-\frac{\delta\theta}{4\pi}\right)\sqrt{r_{20}^2-r_{10}^2}-\left(\frac{\delta\theta}{4\pi}\right)^2\right]=0 \tag{34}$$

6. DETERMINATION OF MAXIMUM WRAPPING AND UNWRAPPING ANGLES:

It can be seen from equation (13) and Figure 3 that the minimum value of r, r_{10}, will yield the maximum value of θ ; i.e.

$$\theta_{MAX.}=\frac{2\pi}{\delta}(r_0-r_{10}) \tag{35}$$

The maximum wrapping angle is found by noting that

$$L=\int_0^{\alpha_{MAX.}}r_2\,d\alpha+\sqrt{r_{2MAX.}^2-r_{10}^2}=\left(r_{20}-\frac{\delta\alpha_{MAX}}{4\pi}\right)\alpha_{MAX}+\sqrt{\left(r_{20}-\frac{\delta\alpha_{MAX}}{2\pi}\right)^2-r_{10}^2}$$

Rearranging terms and squaring the radical leads to $(\alpha_{MAX}\equiv\alpha_M)$

$$\frac{\delta^2}{4\pi^2}\alpha_M^4-\frac{\delta}{2\pi}r_{20}\alpha_M^3+\left(r_{20}^2-\frac{\delta L}{2\pi}-\frac{\delta^2}{4\pi^2}\right)\alpha_M^2-2\left(L-\frac{\delta}{2\pi}\right)r_{20}\alpha_M-r_{20}^2+r_{10}^2+L^2=0 \tag{36}$$

Neglecting terms containing δ , a first approximation for equation (36) produces

$$\alpha_M^2-\frac{2L}{r_{20}}\alpha_M+\frac{1}{r_{20}}\left(L^2+r_{10}^2-r_{20}^2\right)=0 \tag{37}$$

A final approximation can be made if $L^2\gg r_{20}^2\gg r_{10}^2$. That is,

$$\alpha_M^2-\frac{2L}{r_{20}}\alpha_M+\frac{L^2}{r_{20}^2}=\left(\alpha_M-\frac{L}{r_{20}}\right)^2=0\ \Rightarrow\ \alpha_M\simeq\frac{L}{r_{20}} \tag{38}$$

7. DETERMINATION OF r_0, THE INITIAL OUTSIDE RADIUS

If L, δ, r_{10} and r_{20} are given, r_0 is obtained by first noting that the coil length, L, can also be determined from

$$L=\int_0^{\theta_{MAX.}}r\,d\theta+\sqrt{r_{20}^2-r_0^2}=\frac{\pi}{\delta}\left(r_0^2-r_{10}^2\right)+\sqrt{r_{20}^2-r_0^2}$$

Rearranging terms and squaring the radical then yields

$$r_0^4+\frac{\delta^2}{4\pi^2}\left(1-\frac{8\pi L}{\delta}-\frac{8\pi^2}{\delta^2}r_{i0}^2\right)r_0^2+\frac{\delta^2}{\pi^2}\left(L^2-r_{s0}^2+\frac{2\pi L}{\delta}r_{i0}^2+\frac{\pi^2}{\delta^2}r_{i0}^4\right)=0 \tag{39}$$

To obtain a real solution to equation (39), it is required that

$$L \le \frac{\pi}{\delta}\left(r_{s0}^2-r_{i0}^2\right)+\frac{\delta}{4\pi} \tag{40}$$

which is obtained by imposing the condition that the radical term in

$$r_0^2=\frac{\delta^2}{8\pi^2}\left(\frac{8\pi^2}{\delta^2}r_{i0}^2+\frac{8\pi L}{\delta}-1\right)\pm\sqrt{\frac{\delta^4}{64\pi^4}\left(1-\frac{8\pi L}{\delta}-\frac{8\pi^2}{\delta^2}r_{i0}^2\right)^2-\frac{\delta^2}{\pi^2}\left(L^2-r_{s0}^2+\frac{2\pi L}{\delta}r_{i0}^2+\frac{\pi^2}{\delta^2}r_{i0}^4\right)} \tag{41}$$

be real. Taking the positive sign in equation (41) to insure a real root, provides the relation

$$r_0=\sqrt{\left(r_{i0}^2+\frac{\delta L}{\pi}-\frac{\delta^2}{8\pi^2}\right)+\frac{\delta}{\pi}\sqrt{\frac{\delta^2}{4\pi^2}+\left(r_{s0}^2-r_{i0}^2\right)-\frac{\delta L}{\pi}}} \tag{42}$$

8. ALGEBRAIC SIMPLIFICATION OF THE EQUATION OF MOTION (Equation (30))

For the following conditions:

$$r_{s0} > r_0 \quad ; \quad 1 \gg \frac{\delta\theta}{2\pi r_0} \quad ,$$

equation (33) provides the relationship

$$\alpha=\frac{2\pi r_{s0}}{\delta}-\sqrt{\left(\frac{2\pi r_{s0}}{\delta}\right)^2-\left(\frac{4\pi K}{\delta}-\theta\right)\theta} \tag{43}$$

Substituting equation (43) into equation (22) results in

$$r_c=\sqrt{r_{s0}^2-r_0^2} \tag{44}$$

Thus, using equation (44) and taking the time derivative of equation (22)

$$\frac{dr_2}{dt} = \frac{1}{2}(r_2^2 - r^2)^{-\frac{1}{2}}(2r_2\dot{r}_2 - 2r\dot{r}) = 0 \tag{45}$$

By differentiating equations (13) and (14) the following time derivative relations are obtained:

$$\dot{r} = -\frac{\delta}{2\pi}\dot{\theta} \quad , \quad \ddot{r} = -\frac{\delta}{2\pi}\ddot{\theta}$$
$$\dot{r}_2 = -\frac{\delta}{2\pi}\dot{\alpha} \quad , \quad \ddot{r}_2 = -\frac{\delta}{2\pi}\ddot{\alpha} \tag{46}$$

Hence, from equations (45) and (46),

$$r_2\dot{\alpha} - r\dot{\theta} = 0 \tag{47}$$

and,

$$\dot{\alpha} = \frac{r}{r_2}\dot{\theta} \quad ; \quad \ddot{\alpha} = \frac{r}{r_2}\ddot{\theta} - \frac{\delta}{2\pi}\frac{r_c^2}{r_2^3}\dot{\theta}^2 \tag{48}$$

Substituting equations (46), (47) and (48) back into equation (30), simplifying and rearranging yields

$$\ddot{\theta} + A_2\dot{\theta}^2 - A_3\dot{\theta} + A_4 = 0 \tag{49}$$

where,

$$A_2 \equiv \frac{A_0}{A_1 r_2^2}\left[r^2 r_c^2 r_2 + \frac{\delta r r_c^2}{4\pi}(2r_c - r_2) - \frac{\delta^2 r_c^4}{4\pi^2 r_2^3}\right]$$
$$A_3 \equiv \frac{A_0}{A_1} r_c \omega_0 \left(2 r r_c + \frac{\delta r^2}{2\pi r_2} - \frac{\delta r_c}{\pi}\right)$$
$$A_4 \equiv -M/A_1 I + A_0 (r_c \omega_0)^2 r_2 / A_1 \tag{50}$$
$$A_1 \equiv 1 - r_c A_0 (r^2 - r r_2 - \delta r_c / 2\pi)$$
$$A_0 \equiv 8/\pi r_2 r^3$$

These equations together with the initial valves,

$$\theta(t=0) = 0 \quad ; \quad \dot{\theta}(t=0) = 0 \tag{51}$$

completely define the unwrapping angle as a function of time.

9. EXPERIMENTAL PARAMETERS

To evaluate equation (49) for its accuracy in predicting the arming time of an unwinder fuze, a comparison is made between Alfriend's experimental test data and the analytic results. The parameters for Alfriend's experiments are given in Table 1.

10. RESULTS, CONCLUSIONS AND DISCUSSION

Analytic and experimental results are plotted in Figures 4 through 10. Surveying these results, it is found that springs 1, 3, 4, 6 and 7 provide a good correspondence between theory and experiment, with the correlation becoming increasingly better as the angular velocity of the outer case increases. The poorest theoretical-experimental correspondence occurs with springs 2 and 5, although the trend of the relative turns versus time data with increasing outer case spin is predicted.

It is interesting to note that the fall-off of experimental data points from analytically determined points in spring 5 (at a relatively high outer case spin) appears similar to that for the lowest outer case spin rate for spring 1.

Without a thorough knowledge of experimental procedures, devices and data, the reasons for differences between experimental and analytic results cannot rationally be examined. How well the mathematical model will describe the results of an experiment depends on how well the experimental set-up is true to the conditions of the mathematical model. Since the data used in this report must be taken and used as reported, the reason(s) for the discrepencies between analytic and experimental results was not sought.

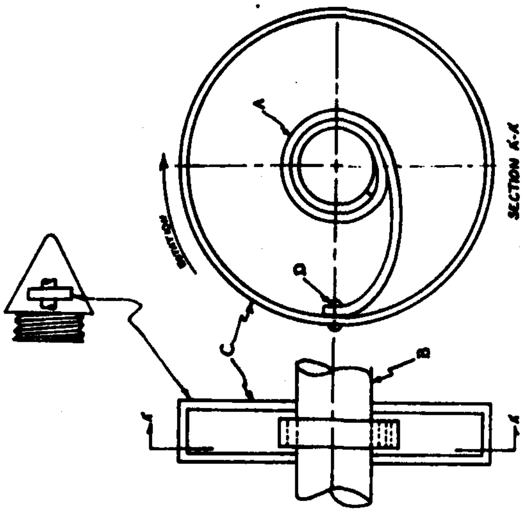


FIGURE 1: DELAYED ARMING DEVICE

Spring No.	Mat'l	E (PSI)	δ (IN)	γ_{10} (IN)	γ_{20} (IN)	b (IN)	L (IN)	ρ_{sw} (LB/IN^3)	M (IN-LB)	θ MAX (RADS)	α MAX (RADS)	ω_0 (RPM)
1	302/ 301 Stainless Steel	$30X10^6$	0.003	0.26	1.0	0.5	71.0	0.29	0.03554	226	71.9	4,200 4,840 5,400 6,000
2	1095 Carbon Steel	$30X10^6$	0.002	0.26	1.0	0.25	38	0.29	0.00526	138	37	4,000 5,000 6,120
3	1095 Carbon Steel	$30X10^6$	0.002	0.4	1.1	0.25	70	0.29	0.00225	157	64.6	3,960 4,800 6,000
4	1095 Carbon Steel	$30X10^6$	0.002	0.5	1.2	0.25	80	0.29	0.0017	157	67.6	3,000 3,711 4,948
5	1095 Carbon Steel	$30X10^6$	0.004	0.26	1.0	0.25	39.5	0.29	0.04212	133.5	40.46	7,000
6	1095 Carbon Steel	$30X10^6$	0.004	0.4	1.1	0.25	70	0.29	0.0202	157	64.6	4,140 4,800 6,666
7	1095 Carbon Steel	$30X10^6$	0.004	0.5	1.2	0.25	80	0.29	0.0136	146.1	67.6	3,261 3,711 4,999

TABLE 1: Experimental Parameters From "Study of Wrapped Springs for Application to a Delayed Arming Mechanism" - Summary Report ER-1404-Aircraft Armaments Inc. Contract No. DA-36-034-501-ORD-74RD, June 1958

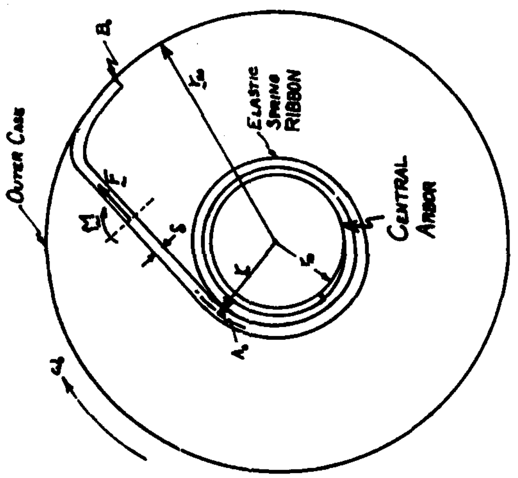

Outer Case
Elastic Spring Ribbon
Central Arbor
B
A
M
F
S
ω

FIGURE 2: T.B. ALFRIEND'S PROBLEM

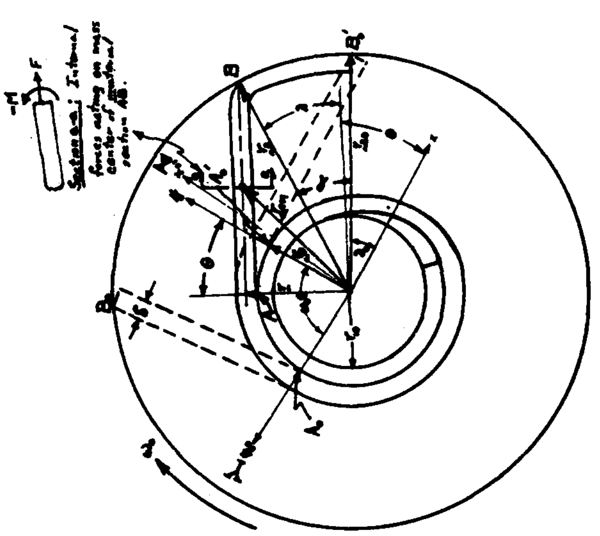

-M
F
Section a-a: Internal forces acting on mass center of sectional section AB.
B
B'
A
θ
ω

FIGURE 3: GEOMETRY, TRACTIONS AND POSITIONAL IDENTIFIERS FOR MODIFIED ALFRIEND PROBLEM

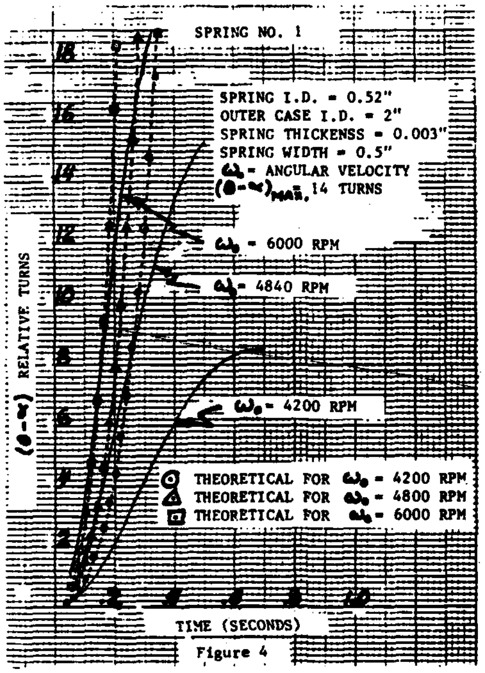


Figure 4

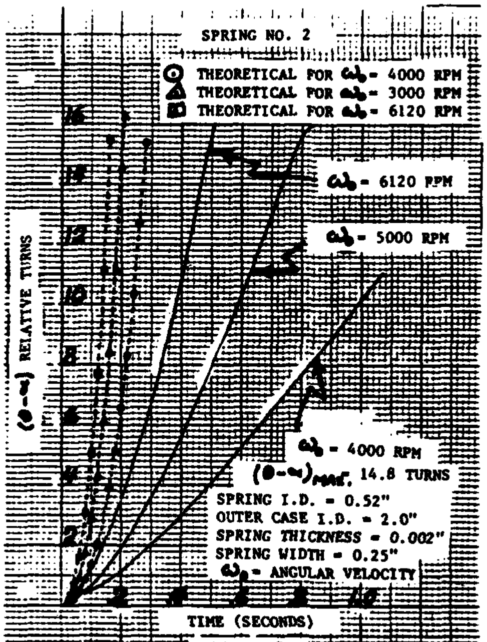


Figure 5

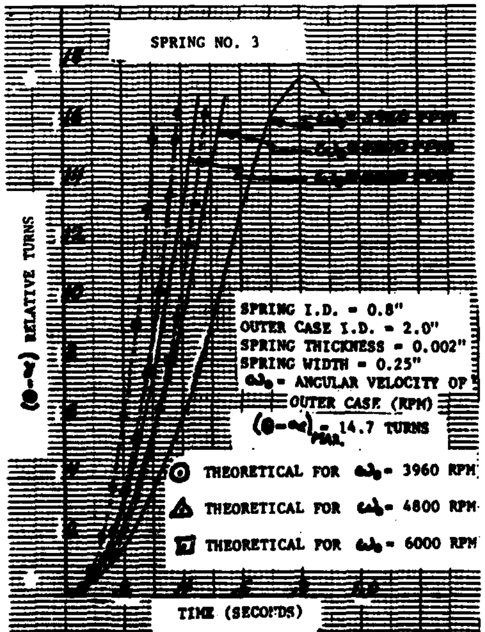


Figure 6

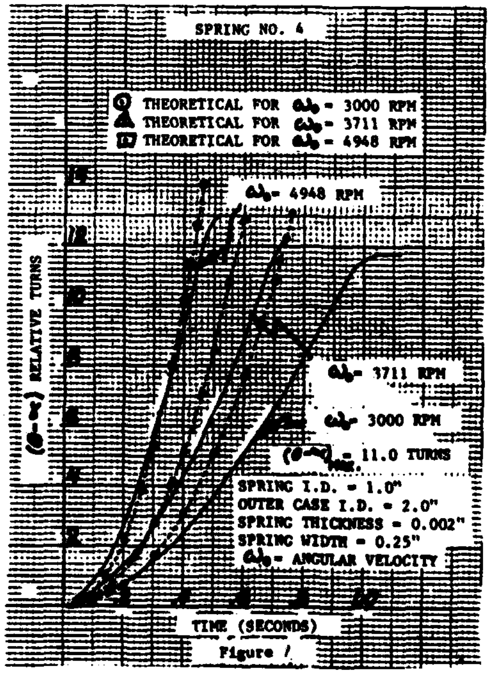


Figure 7

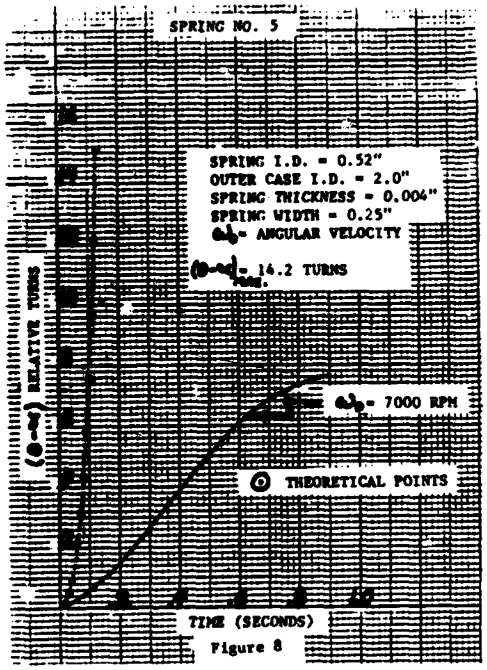


Figure 8

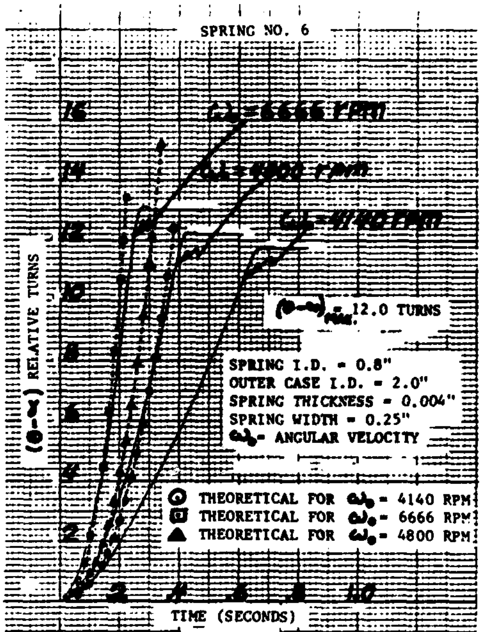


Figure 9

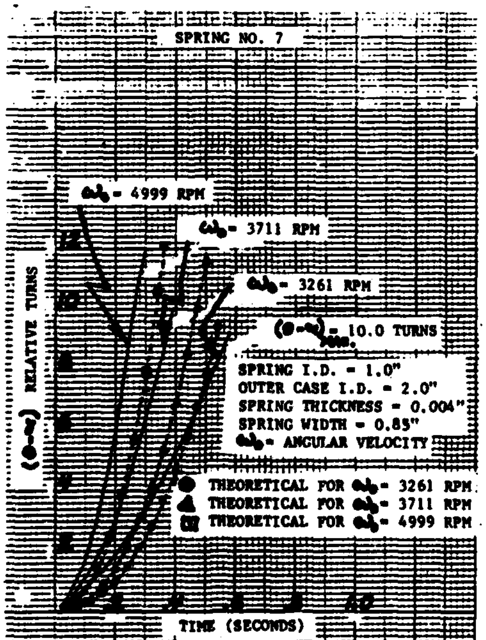


Figure 10

PHOTONIC SEEKER DEVELOPMENT (U)

*JOSEPH G. DUTHIE, DR., PAUL R. ASHLEY, DR.
RESEARCH DIRECTORATE, US ARMY MISSILE LABORATORY
US ARMY MISSILE COMMAND
REDSTONE ARSENAL, ALABAMA 35898
AND
JURIS UPATNIEKS, MR.** AND H. K. LIU, DR.***

INTRODUCTION

Ideally, a seeker system to be used for target identification and missile guidance should be able to recognize the target over an extended set of ranges, orientations and aspects. In the most simple scenario a single feature such as a characteristic infrared emission or designation of the target by an aimed laser spot will suffice. However, for autonomous seekers which have to seek out and destroy energy targets without the aid of laser designators or where the differentiation between friend or foe is more difficult, far more information has to be processed than may be possible by electronic means.

Consider the problem of identifying and correctly locating a single scale, orientation and aspect of an enemy vehicle in a snapshot taken with a typical instamatic camera. For this example we will take the focal length of the lens to be 2 cm, the lens aperture to be 1 cm, the wavelength of the light to be 5 x 10^{-5} cm and the field of view 1 cm x 1 cm. The numbers are useful only for order of magnitude calculation and need not represent an actual system. In the image plane the size of the resolution elements is 10^{-4} cm. Thus, the image may be thought of as a 10^4 x 10^4 array of resolvable elements. To recognize and locate the tank, the optimum estimator can be shown to be the cross correlation between a known image of the tank and the image to be searched. For an image consisting of 10^4 x 10^4 resolvable points, this involves computing two separate 10^4 x 10^4 point Fourier Transforms, multiplying the two together then forming the 10^4 x 10^4 point Fourier Transform of the resultant. A total of about 10^{10} multiplications is involved. The task of doing this at T.V. frame rates requires

---

** Permanent Address: Environmental Research Institute of Michigan
P.O. Box 8618, Ann Arbor, MI 48107

*** Permanent Address: University of Alabama, Tuscaloosa, AL 35486 about 10^{12} multiplications per second and is well beyond the capability of any digital computer. Of course one could relax the requirements considerably in terms of number of resolution elements or processing rate, but only at the cost of reduced performance.

The photonic correlator, however, has such a capability today. It has the added advantages of low power consumption, mechanical simplicity and it can be made small enough to fit comfortably into a submissile. There are, however, several limitations which must be overcome before the use of photonic correlators become a reality for missile systems. These include the cost of the light modulator and the need for extending the memory of the photonic correlator to include an extended set of target scales, sizes and orientations. The problem of cost for the light modular appears to be one of manufacturing techniques and it should be possible to produce low cost devices cheaply enough in suitable quantities. The need to extend the memory of the correlator has been the subject of much of our recent research on the use of photonic correlators.

The MICOM Photonic Correlator - An Overview

The science of photonic optical information processing relies heavily on the Fourier Transforming Properties of Lenses (1). This was applied by A. B. Vander Lugt of the University of Michigan's Radar Laboratory, who, in 1963, demonstrated a new technique for synthesizing matched filters for coherent processors (2) (3).

Figure 1 represents the two stages of coherent image processing using Vander Lugt filtering techniques. In Figure 1A, a reference scene on a transparency of amplitude transmittance (x,y) is illuminated by a source of coherent light. The transparency is in the front focal plane of the Fourier Transforming lens L_1. In the back focal plane is located a photographic plate. Exposure of the plate to the transformed image (x,y) simultaneously with a reference beam produces what is called a matched spatial filter of the reference scene. Subsequent to development, the plate is re-inserted in the optical system as in Figure 1B. In this case the reference beam is removed and a new input transparency g(x,y) is inserted. The property of the matched filter is that in the arrangement shown in Figure 1B, the detected output in the cross correlation of f(x,y) and g(x,y), i.e., a bright spot will appear in the output plane if g(x,y) contains sufficient information about f(x,y). Furthermore, the position of the spot is a function of the position of the object located in g(x,y). Thus, the photonic correlator can effectively detect the presence of a test object in a given field of view.

As an example of the technique Figure 2 shows an image of a sedan used to make a Vander Lugt filter for the photonic correlator. Figure 3 shows the response of the correlator to a series of scenes showing the same vehicle driving along a highway. The lower set of photographs show the response of the photonic correlator to the input scenes. A bright correlation spot appears in the output. The existence of this spot identifies the existence of the automobile in the input scene and the position of the spot denotes the position of the automobile. In a dynamic situation the photonic correlator follows the target at T.V. frame rates.

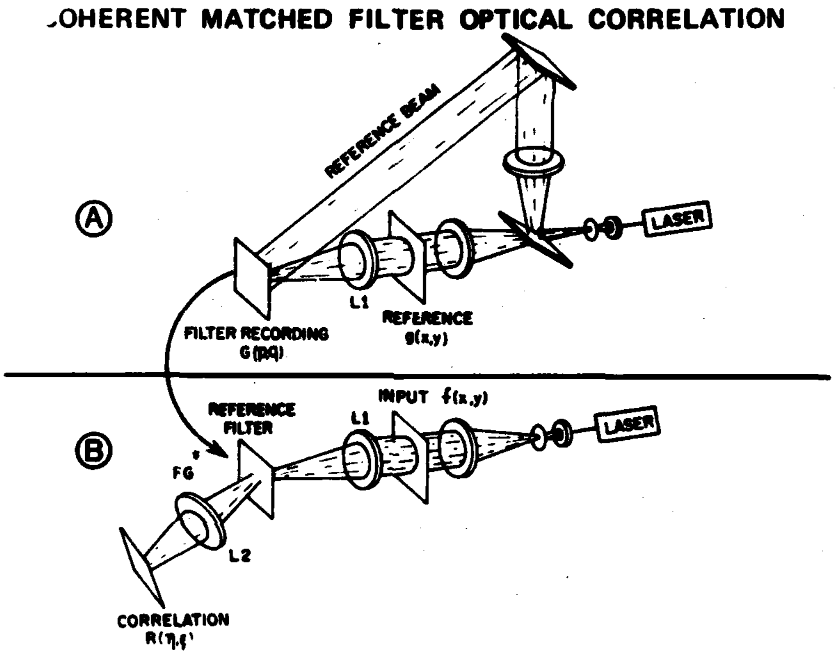

COHERENT MATCHED FILTER OPTICAL CORRELATION
Ⓐ
REFERENCE BEAM
LASER
L1
REFERENCE
g(x,y)
FILTER RECORDING
G(p,q)
Ⓑ
REFERENCE
FILTER
INPUT f(x,y)
L1
LASER
FG*
L2
CORRELATION
R(η,ξ)

Figure 1. Typical Optical Correlator (A) Method of Recording Filters, (B) Method of Obtaining Correlations.

Real-time applications of the photonic correlator were demonstrated by Guenther et. al (4) of MICOM by utilizing a liquid crystal light modulator to generate the spatially coherent input scene. This enabled the correlator to recognize and track targets at about T.V. frame rates. In 1980, Duthie et. al (5) also of MICOM reported a further improvement in the photonic correlator, namely, the use of solid state laser diodes rather than bulky, fragile gas lasers. The combination of the use of real-time input devices and solid state light sources spurred interest in developing a compact correlator as a practical tracking seeker. Figure 4 shows a possible configuration of such a system.

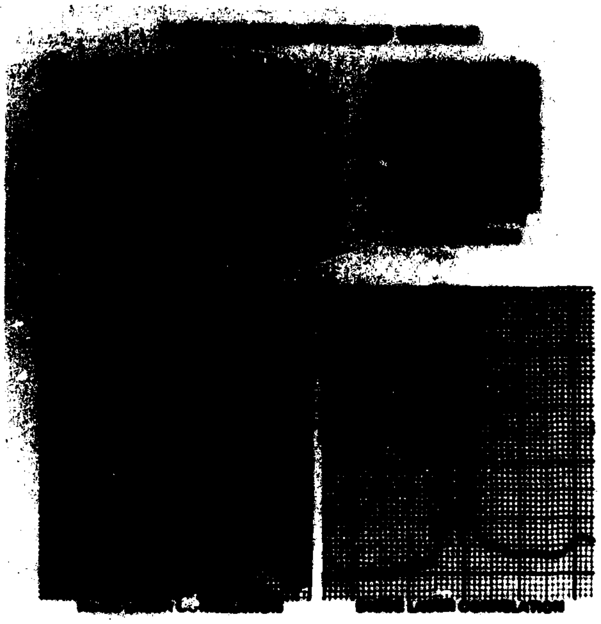


Figure 2. Example of Correlator Input (a) Input Object (b) Matched Filter of some Object.

Increased Memory for the Photonic Correlator

Major improvements need to be made in the capacity of a photonic correlator to store and address a large number of reference images before the photonic correlator can be regarded as a serious contender for missile guidance and tactical homing applications. Leib et. al (6) have designed an optical matched filter correlation system in which a large array of holograms can be stored in a matched filter through the use of a multiple number of holographic lenses. Theoretical predictions of a capability to store up to 2500 matched filters have yet to be realized in the laboratory.

A new method of addressing an array of matched filters has been developed by Liu and Duthie (7) at MICOM. In the basic correlator system of Figure 1, a phase screen is placed in the input plane. This screen has high optical transparency but spatially modulates the phase of the light waves. The periodic modulation is chosen so as to generate an array of islands in the Fourier Transform plane of the correlator. Figure 5 shows a photomicrograph of the phase screen while Figure 6 shows the array of islands found in the transform plane. Thus, the effect of the phase screen is to modify the light distribution in the back focal plane of the transform lens so that instead of a single spot, an array of spots is produced by a collimated input beam. In the present case, it has been found that a 5 x 5 array of spots can be engineered to carry the bulk of the light intensity.

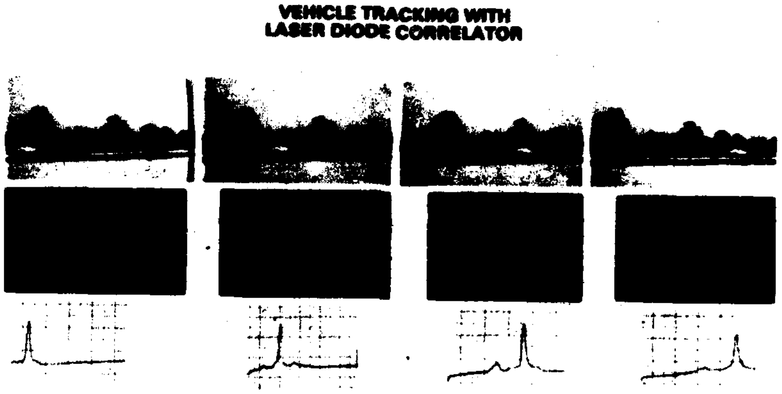

VEHICLE TRACKING WITH LASER DIODE CORRELATOR

Figure 3. Tracking of Automobile using Optical Correlator

Furthermore, Liu and Duthie have demonstrated that at each spot, a matched filter, of the Vander Lugt type, can be located. These workers have demonstrated the effectiveness of each of the 25 spots individually at target identification and tracking.

Table 1 shows the measured intensities of the spots in the 5 x 5 array. The central (0,0) spot is certainly the brightest. The remainder, although not strictly equal are indeed equal to within an order of magnitude. Autocorrelations have been measured for matching input scenes with recorded filters of the same scene for each of the spots addressed individually. Previous work at MICOM has indicated that up to eight matched filters can be recorded at every spot or island in the Transform Plane. Thus, the potential exists for the storage of up to 200 matched filters in this system.

order	-2	-1	0	+1	+2
+2	.35	.35	1.17	.32	.23
+1	.43	.46	1.43	.42	.39
0	.96	.98	2.92	1.01	.91
-1	.48	.40	1.39	.41	.37
-2	.28	.29	1.00	.28	.27

Table 1. The Measured Relative Intensity Values in μW.

To date we have not made a complete filter array and addressed each in parallel. This requires precise alignment in the filter manufacturing stage and has to be done under computer control. Such an effort is currently underway at MICOM and should be completed shortly.

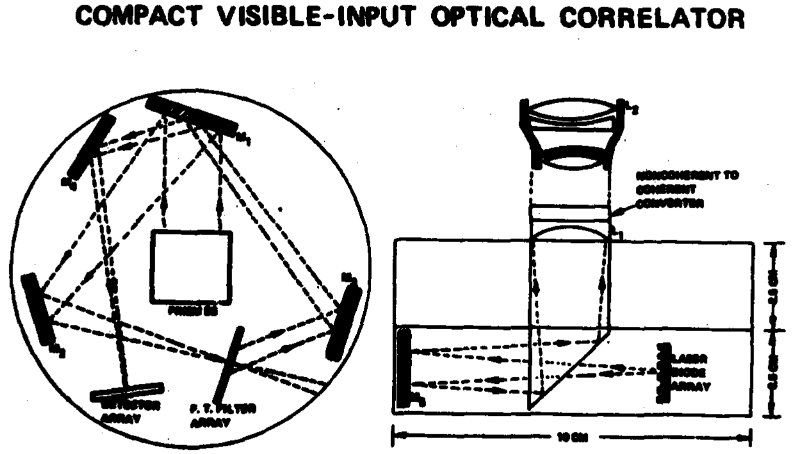

COMPACT VISIBLE-INPUT OPTICAL CORRELATOR
DETECTOR ARRAY
F. T. FILTER ARRAY
NONCOHERENT TO COHERENT CONVERTER
LASER DIODE ARRAY
10 CM

Figure 4. Proposed Design for Compact Optical Correlator.

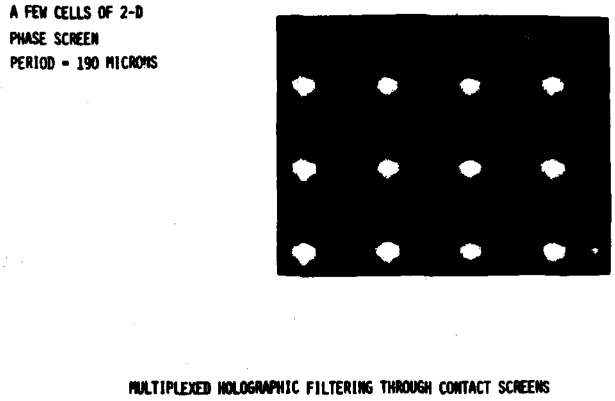


Figure 5. Photomicrograph of Phase Screen used to Multiplex the Number of Filters Addressed in Parallel. Period = 1/133 inches.

Non-Coherent Illumination

Whereas the optical correlator relies on the Fourier Transforming Properties of Lenses and requires good spatial coherence of the illuminating beam, no requirements are needed for temporal coherence of the light source. Indeed, several advantages may accrue if the laser or laser diode in the photonic correlator could be replaced by a thermal light source. First would be the advantage of cost, followed by advantage in terms of the absence of coherent or artiface noise. This noise is a direct consequence of the temporal and spatial coherence of the source and can be eliminated by using a broad band light source.

There is, however, a major problem in using a thermal source. The Fourier Transforming Properties of Lenses is a function of the wavelength of the light used. Thus, if a broad spectral source is used to read out a matched filter, then it will only be able to form a correlation over a very narrow band of wavelengths close to the wavelength used in making the filter. For example, Figure 7 shows the Fourier Transform of a

2-Dimensional Ronchi Grating when viewed in a coherent image processing system when illuminated by white light.

Instead of a set of discrete spots on a rectangular array, the lens generates a set of rainbow colored lines showing extreme chromaticity in the Fourier Transforming function

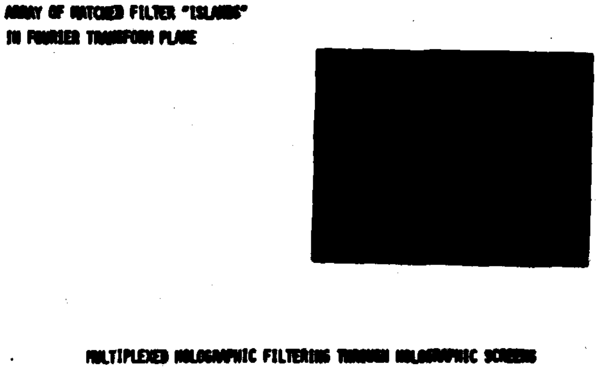

ARRAY OF MATCHED FILTER "ISLANDS"
IN FOURIER TRANSFORM PLANE
MULTIPLEXED HOLOGRAPHIC FILTERING THROUGH HOLOGRAPHIC SCREENS

Figure 6. Array of Spots or Islands in the Transform Plane Produced by the Phase Screen, At each Spot in the Central 5 x 5 Region up to Eight Matched Filters can be Stored and Addressed in Parallel.

A solution to this problem has been developed at the University of Rochester (8) and has been improved on by Duthie and Upatnieks at MICOM. The optical arrangement is shown in Figure 8. The transforming process of a simple lens has been replaced by a train of two off axis holographic lenses, a simple refractive lens and a diffraction grating. The first lens is a combination of a thin lens of focal length F_0 and an off axis holographic lens of focal length $- F_0$ at a wavelength λ_0. At other wavelengths it has a focal length $- F_0 \lambda_0/\lambda$. At λ_0 the effect of this lens is to diffract an input collimated into a collimated beam at an angle to the optical axis. At larger and shorter wavelengths the diffracted beams diverge or

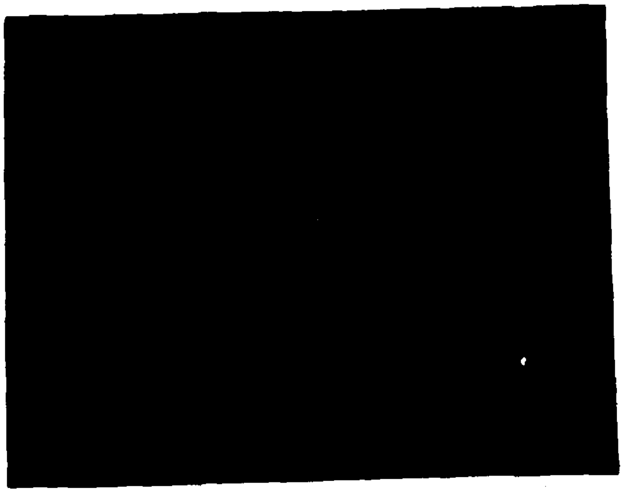


Figure 7. Conventional White Light Fourier Transform of a 2-Dimensional Ronchi Grating. Note the Extended Nature of the Individual Orders due to Severe Chromatic Effects.

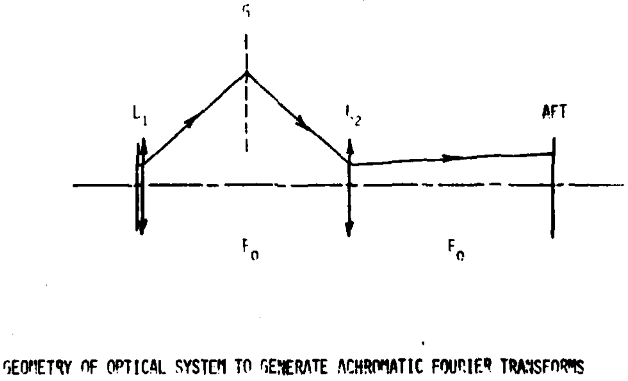

G
L1
L2
AFT
F0
F0
GEOMETRY OF OPTICAL SYSTEM TO GENERATE ACHROMATIC FOURIER TRANSFORMS

Figure 8. Optical Arrangement for Achromatic Fourier Transform System.

converge according to wavelength. The grating serves to deflect the beams back onto the converging holographic lens L_2 which has a focal length $+ F_0 \lambda_0/\lambda$. The net effect of these diffractions and propagations of the waves is to generate on the focal plane a Fourier Transform which is, to a high approximation, independent of wavelength. Figure 9 shows the Fourier Transform in white light obtained of the same 2-Dimensional Ronchi Grating as used for Figure 7. The chromaticity in the Fourier Transform plane has been effectively removed.

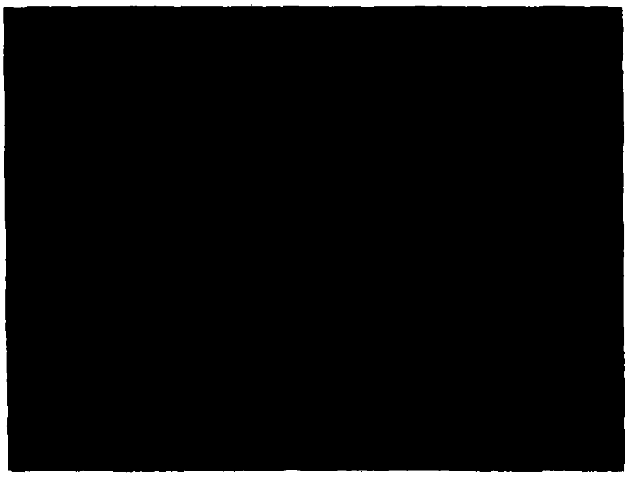


Figure 9. White Light Fourier Transform of 2-Dimensional Ronchi Grating using Optical Arrangement of Figure 8.

This type of optical arrangement has been effectively used to generate auto correlations between input and identical reference scenes. An additional element needs to be added to the system to make the final correlation achromatic - in this case a simple diffraction grating was used.

Results so far have shown a dramatic improvement in the apparent signal to noise of correlations obtained using white light rather than laser light in the achromatic correlator. Apart from the data and noise in the input scene, none of the other noise in the system was transformed achromatically and thus did not correlate over the entire spectrum of the source used - in this case a high intensity Hg vapor lamp.

Conclusions

Experiments have been performed which indicate a means to extend significantly the number of matched filters which can be simultaneously addressed in the photonic correlator. Results indicate that a total of 200 independent images of the target can be interrogated in parallel.

This corresponds to an effective computation of over 10^{13} complex multiplications per second. The suggested method is mechanically simple, optically easy to implement, requires little power and can be fabricated into a compact unit.

Improvement in the signal to noise of a photonic correlator can be achieved by using an achromatic transform system and a thermal light source. At this time, however, the throughput of such a system is not sufficient to make a practical device. Whereas, efforts should continue to develop techniques to correlate using thermal light sources, the principal thrust for making a field operational system should, at this time, be concentrated on the use of laser diode sources together with the use of phase screens to extend the memory of these systems.

REFERENCES

1. See for Example, J. W. Goodman, "Introduction to Fourier Optics", McGraw-Hill, (1968).

2. A. B. Vander Lugt, "Signal Detection by Complex Spatial Filtering", Radar Lab Report 4594-22-T, (1963).

3. A. B. Vander Lugt, "Signal Detection by Complex Spatial Filtering", IEEE Trans. Inform. Theory, IT-10:2, (1964).

4. B. D. Guenther, C. R. Christensen and J. Upatnieks, "Coherent Optical Processing: Another Approach", IEEE J. Quantum Electronics, WE-15, 1348, (1979).

5. J. G. Duthie, J. Upatnieks, C. R. Christensen, and R. D. McKenzie, "Real Time Optical Correlation with Solid State Sources", SPIE 231, 281, (1980).

6. K. G. Leib, R. A. Bondurant, M. R. Wohlers, "Optical Matched Filter Correlator Memory Techniques and Storage Capacity", Opt. Engr. 9, 414, (1980).

7. H. K. Liu and J. G. Duthie, "Real Time Screen Aided Multiple Image Optical Holographic Matched Filter Correlator", submitted for publication, Applied Optics, 1982

8. G. M. Morris, "Diffraction Theory from Achromatic Fourier Transformation", Appl. Opt. 11, 1255 (1972).

TITLE: STATIONARY AND TRAVELING LOADS IN A HOLLOW CYLINDER

ALEXANDER S. ELDER, MR.
JAMES N. WALBERT, DR.
KATHLEEN L. ZIMMERMAN, MRS.
US ARMY BALLISTIC RESEARCH LABORATORY
ABERDEEN PROVING GROUND, MD 21005

I. INTRODUCTION

Just after World War II Dr. R. Beuwkes [1] of Watertown Arsenal introduced the senior author to the theory of elastic stresses in thick-walled cylinders and the technology of shell-pushing tests. The research at Watertown Arsenal culminated in the publication of the Thick Walled Cylinder Handbook [2], a monument to skill in analysis and computations on a desk calculator. Investigation of non-metallic rotating bands at BRL led to renewed interest in this area. We found that we were unable to interpolate in the tables cited above due to their limited accuracy; moreover the value of Poisson's ratio used in the computations was not appropriate for modern gun steels. An independent investigation was initiated, using residue theory in place of Fourier series [3,4,5]. Since the eigenvalues were complex, we required a subroutine for Bessel functions of integral order and complex argument. The required subroutine was developed at BRL. A Gauss continued fraction was used to reduce round off errors inherent in series calculations [6,7]. A code giving accurate stresses on the outside of the gun tube was developed [8,9]. This code was used to calculate strains in a highly instrumented gun tube. The appropriate value of Poisson's ratio and Young's modulus was obtained from the Benet Laboratory. Agreement between theory and measurement was good at low velocities, but systematic deviations were observed at high velocities. This result was forecast in an early paper by G.S. Taylor [10], who used a dynamic version of the Winkler theory for thin-walled tubes, but his results were apparently ignored by the Army. A program based on scalar and vector wave functions was initiated at BRL. The computations are difficult except for torsion, which we discuss below. The theoretical work shows that the equilibrium stress distribution is obtained when the velocity of travel approaches zero in the limit, as one would expect on physical grounds.

Recently we have resolved difficulties in calculating stresses near a discontinuity of loading on the inner surface of the cylinder. A method of calculating elastic stresses in thin-walled cylinders was also derived, so we are able to use the same mathematical formulation for wall ratios ranging from .01 to 5. Both of these problems required asymptotic methods and involved large values of the complex transform variable in the analysis.

II. NUMERICAL DIFFICULTIES

Formulation of boundary value problems for the infinite hollow cylinder has followed traditional lines and is not exceptionally difficult. Real problems arise in the numerical evaluation of Fourier integrals and the generation of Bessel functions of the second kind due to the integer arithmetic of the digital computer and its limited exponent range. Memory requirements and execution time are relatively modest for the class of problems under consideration. We have considered four types of error in the course of programming and numerical analysis.

Round off error is persistent and insidious. It is very severe in the evaluation of Fourier integrals by quadratures along the real axis and was the principle reason why the calculus of residues was used in preference. It occurred in acute form in calculating Bessel functions of the second kind. This difficulty motivated our development of the subroutine cited above. In this paper we discuss round off error occurring in the manipulation of asymptotic series. Round off error is also a principle concern in generating special functions by recursion formulas, where it arises in connection with stability criteria.

A continued fraction obviously can be used only for values of the variable and parameter for which division by zero will not occur. We finally are able to prove that division by zero would not occur in the portion of the subroutine using Gauss continued fractions. Theorems of Bucholz [11] and Hurwitz [12] were required in the proof [13]. The analysis is closely related to Hurwitz stability theory.

Serious truncation error has occurred only in evaluating residue series for the inner radius at points very close to the discontinuity of loading. Only recently have we found a method for improving the convergence of the residue series.

III. STRESSES NEAR A DISCONTINUITY OF LOADING FOR AXIALLY SYMMETIC STRESSES

For brevity we consider only axial stresses produced by a step function of pressure or shear applied to the inner cylindrical surface. The analysis of tangential stresses is similar. We superimpose a constant stress and a discontinuity stress to obtain the step function. For pressure loading, we have

$$\tau_{rz} = 0, \quad \sigma_r = \tfrac{1}{2}\,\sigma_0, \quad r = a \quad \text{and} \tag{1a}$$

$$\tau_{rz} = 0, \quad \sigma_r = -\tfrac{1}{2}\,\sigma_0, \; z<0; \; \sigma_r = \tfrac{1}{2}\,\sigma_0, \; z>0, \; r = a, \tag{1b}$$

respectively, and for shear loading

$$\sigma_r = 0, \quad \sigma_{rz} = \tfrac{1}{2}\,\tau_0, \quad r = a \tag{2a}$$

$$\sigma_{rz} = 0, \quad \tau_{rz} = -\tfrac{1}{2}\tau_0, \; z<0; \; \tau_{rz} = \tfrac{1}{2}\tau 0, \; z>0, \; r = a \tag{2b}$$

In both cases

$$\sigma_r = 0, \quad \tau_{rz} = 0, \quad r = b. \tag{3}$$

The solutions corresponding to (1a) and (2a) can be obtained by elementary methods and will not be considered here. The discontinuous stresses in (1b) and (2b) are represented by Cauchy discontinuous factors to facilitate solution by separation of variables.

The stresses are derived from Love's stress function [14] in the form

$$\phi_c = [AI_0(sr) + BK_0(sr) + CsrI_1(sr) + DsrK_1(sr)]\cos(sz) \tag{4}$$

for pressure loading and

$$\phi_s = [AI_0(sr) + BK_0(sr) + CsrI_1(sr) + DsrK_1(sr)]\sin(sz) \tag{5}$$

for shear loading. If the boundary conditions are homogeneous, we obtain four homogenous linear equations which are satisfied only if the determinant of the coefficients is equal to zero. In the case of shear loading, we obtain from Eq. (1) and a number of intermediate calculations the characteristic equation

$$\Delta_c = 0 \tag{6}$$

where

$$\Delta_c(s) = \begin{vmatrix} I_0(p) & K_0(p) & \alpha_1 I_1(p) & \alpha_1 K_1(p) \\ I_1(p) & -K_1(p) & pI_0(p) & -pK_0(p) \\ I_0(q) & K_0(q) & \beta_1 I_1(q) & \beta_1 K_1(q) \\ I_1(q) & -K_1(q) & qI_0(q) & -qK_0(q) \end{vmatrix} \tag{7}$$

and

$$p = sa,\ \alpha_1 = [p + (2 - 2\nu)/p],\quad q = sb,\ \beta_1 = [q + (2 - 2\nu)/q]. \tag{8}$$

The shear loading leads to the characteristic equation

$$\Delta_s(s) = 0 \tag{9}$$

where

$$\Delta_s(s) = - \Delta_c(s) \tag{10}$$

and obviously has the same characteristic roots.

The characteristic roots in the first quadrant of the complex s plane have the approximate value

$$s_n = t_n/(b-a) \tag{11}$$

where

$$t_n \sim \log_e[(2n-1)\pi] + i\ (n-\tfrac{1}{2})\pi,\ n>1 \tag{12}$$

The approximate values of s_n obtained from Eq. (11) are improved by Newton's method in the complex plane. It should be observed that all the determinants occurring in the analysis are analytic functions of s even though logarithms occur in the series for the modified Bessel function of the second kind.

We find that

$$\sigma_z = \frac{\sigma_0}{\pi} \int_0^{\infty} (\Delta_1/\Delta_c)\sin(sz)ds/s \tag{13}$$

is the solution corresponding to pressure loading, Eq. (1b), and

$$\tau_{rz} = \frac{\tau_0}{\pi} \int_0^{\infty} (\Delta_2/\Delta_s)\cos(sz)ds/s \tag{14}$$

is the solution to the shear problem, Eq. (2b). The determinants Δ_1 and Δ_2 are given by

$$\Delta_1 = \begin{vmatrix} 0 & 0 & -2I_0(q) - \beta_2 I_1(q) & 2K_0(q) + \beta_2 K_1(q) \\ I_1(p) & -K_1(p) & pI_0(p) & -pK_0(p) \\ I_0(q) & K_0(q) & \beta_1 I_1(q) & \beta_1 K_1(q) \\ I_1(q) & -K_1(q) & qI_0(q) & -qK_0(q) \end{vmatrix} \tag{15}$$

$$\Delta_2 = \begin{vmatrix} I_0(p) & K_0(p) & 2I_0(p) + pI_1(p) & -2K_0(p) + pK_1(p) \\ I_1(p)/p & -K_1(p)/p & 3I_0(p) - \alpha_2 I_1(p) & -3K_0(p) - \alpha_2 K_1(p) \\ I_0(q) & K_0(q) & \beta_1 I_1(q) & \beta_1 K_1(q) \\ I_1(q) & -K_1(q) & qI_0(q) & -qK_0(q) \end{vmatrix} \tag{16}$$

where

$$\alpha_2 = (2-2\nu)/p, \qquad \beta_2 = (2-2\nu)/q \tag{17}$$

We obtain asymptotic approximations of the integrands by using Wronskian relations connecting $I_0(x)$, $I_1(x)$, $K_0(x)$, and $K_1(x)$, where $x = p$ or $x = q$, and the leading terms of the Hankel asymptotic expansions [15]. The leading terms are

$$I_0(x) \sim \frac{e^x}{(2\pi x)^{\frac{1}{2}}} \left[1 + \frac{1^2}{1!8x} + \frac{1^2 \cdot 3^2}{2!(8x)^2}\right] \tag{18}$$

$$I_1(x) \sim \frac{e^x}{(2\pi x)^{\frac{1}{2}}} \left[1 - \frac{1 \cdot 3}{1!8x} - \frac{1^2 \cdot 3 \cdot 5}{2!(8x)^2}\right] \tag{19}$$

$$K_0(x) \sim \frac{\pi}{2x}^{\frac{1}{2}} e^{-x} \left[1 - \frac{1^2}{1!8x} + \frac{1^2 \cdot 3^2}{2!(8x)^2}\right] \tag{20}$$

$$K_1(x) \sim \frac{\pi}{2x}^{\frac{1}{2}} e^{-x} \left[1 + \frac{1 \cdot 3}{1!8x} - \frac{1^2 \cdot 3 \cdot 5}{2!(8x)^2}\right] \tag{21}$$

We find

$$\Delta_1/(s\Delta_c) = s/(s+s_0), \qquad \Delta_2/(s\Delta_s) = - s/(s+s_0) \tag{22}$$

where

$$s_0 = (7-8\nu)(b-a)/4ab \tag{23}$$

On combining Eqs. (13), (14), and (22) we find the resulting integrals can be expressed in terms of sine and cosine integrals [16]. The approximation for small z follows from the fact that large values of the transform variable s correspond to small values of the argument z, according to the usual theory of Fourier integrals. Let γ be Euler's Constant in this context and let $z_0 = 1/s_0$ be a characteristic length. Then, when z is positive and very small, we have approximately

$$\sigma_z \approx \tfrac{1}{2} \sigma_0 \tag{24}$$

for pressure loading and

$$\tau_{rz} \approx \frac{\tau_0}{\pi} \; [\gamma + \log_e(z/z_0)] \qquad (25)$$

for shear loading. We represent the logarithm by an integral of Fourier type [17, 18]. We subtract these dominant terms from the integrals given in Eqs. [13] and [14]. We obtain

$$\sigma_z = \sigma_0 \left[\tfrac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{\Delta_1 - \Delta_c)\sin(sz)}{s\Delta_c(s)} \, ds \right] \qquad (26)$$

$$\tau_{rz} = \frac{\tau_0}{\pi} \left[\log_e(z/z_0) + \int_0^\infty \frac{\Delta_2 \cos(sz_0) \, ds}{s\Delta_c(s)} \right. \qquad (27)$$

$$\left. + \int_0^\infty \frac{(\tfrac{1}{2} \Delta_2 - \Delta_s)(2\cos(sz) - 2\cos(sz_0) ds}{s \, \Delta_s(s)} \right]$$

The integrals in Eqs. (26) and (27) are more rapidly convergent than the original integrals in Eqs. (13) and (14) and will lead to more rapidly convergent residue series when the limits of integration are taken between $-\infty$ and ∞. The integrals must be re-written in exponential form as illustrated in the torsion problem to insure convergence of the contour integrals.

VI. ELASTIC STRESSES IN THIN WALLED CYLINDERS

We observe from Eqs. (11) and (12) that the eigenvalues of high order for a thin-walled cylinder become very large in absolute value. Exponential over-run then occurs when we use the Hankel asymptotic series to evaluate the various determinants. Moreover, when we use Laplace's reduction of the determinant in Eq. (7), we find expressions like $p[I_0^2(p) - I_1^2(p)$ and $p[K_0^2(p) - K^2(p)]$ occur, together with similar expressions involving q. When these expressions are evaluated by means of the Hankel asymptotic expansions, the leading terms are cancelled by subtraction, leading to increasingly severe round off error as the wall ratio approaches one. To overcome these difficulties, we obtained asymptotic expansions of these expressions in which the subtraction occurs algebraically rather than numerically. The exponentials were also combined algebraically, thus eliminating exponential over run for the range of wall ratios of interest.

Let [19]

$$w = Aw_1 + Bw_2 + Cw_3, \tag{28}$$

where

$$w_1 = \pi[pI_0{}^2(p)-pI_1{}^2(p)] \tag{29}$$

$$w_2 = [pI_0(p)K_0(p)+pI_1(p)K_1(p)] \tag{30}$$

$$w_3 = [pK_0{}^2(p)-pK_1{}^2(p)]/\pi, \tag{31}$$

Then w satisfies the following differential equation.

$$p^3w''' + 2\ p^2w'' - (4p^2+p)w' + w=0 \tag{32}$$

We find

$$w_2 = \sum_0 a_n\ p^{-n}\ , \tag{33}$$

where the odd numbered coefficients are zero, $a_0 = 1$, $a_2 = -\frac{1}{8}$, and

$$a_n=(n^3 - 5n^2 + 7n-3)/(4na_{n-2}) \tag{34}$$

for $n>2$ and even.

We let $w_1 = e^{2p}\ w_1$, $w_3 = e^{-2p}\ w_3$. Then

$$p^3w_1''' + (6p^3 + 2p^2)w_1'' + (8p^3+8p^2-2)w_1' + (8p^2-2p+1)\ w_1 = 0 \tag{35}$$

with a similar equation for w_3. We find

$$W_1 = \sum_{1} b_n p^{-n} \tag{36}$$

where $b_1 = 1$, $b_2 = 1/8$, and

$$(8n-8)\, b_n = (6n^2-14n+6)\, b_{n-1} - (n^3-5n^2+7n) b_{n-2},\ n>2 \tag{37}$$

The function w_3 was treated in a similar manner. These formulas were programmed. We obtained 500 eigenvalues for a series of wall ratios ranging from .01 through 5, and the corresponding stresses at the outside radius, where the residue series is rapidly convergent.

V. STRESSES DUE TO AN ACCELERATING LOAD

We outline a method of analysis based on superposition, an eigenvalue expansion, interchange in the order of integration, and the evaluation of a complicated infinite integral. Justification for the various steps is omitted for brevity, but will be presented elsewhere in due course.

We assume the outside cylindrical surface is free of stress, but the inner boundary is subject to a discontinuous moving load.

$$\tau_{r\theta} = 0, \quad r = b \tag{39}$$

$$\tau_{r\theta} = \tau_0\, F_0(z,t) \tag{39}$$

where

$$F_0(z,t) = \tfrac{1}{2}, \quad z>T(t) \tag{40a}$$

$$= -\tfrac{1}{2}, \quad z<T(t) \tag{40b}$$

and T(t) is the travel. We assume the velocity is subsonic, that is, $\dot{T}(t)<c_2$ where c_2 is the velocity of the shear wave in steel. In order to use separation of variables we assume

$$F_0(z,t) = \frac{1}{\pi}\int_0^\infty \frac{\sin[s\beta(z,t)]ds}{s} \tag{41}$$

where

$$\beta(z,t) = z - T(t) \tag{42}$$

We assume a solution of the form

$$\tau_{r\theta} = F_0(z,t)\ [R_0(r) - \sum_1 Q_n R(q_n,r)] + \sum_1 Q_n R(q_n,r)F_n(z,t) \tag{43}$$

where

$$R_0(r)=[a^2(b^4-r^4)]/[r^2(b^4-a^4)] \tag{44}$$

and

$$R(q_n,r) = A(q_n)[I(q_n r)K_2(q_n b)-K_2(q_n r)Y_2(q_n b)] \tag{45}$$

The eigenfunction expansion

$$R_0(r) = \sum_1 Q_n\ R(q_n,r) \tag{46}$$

can be obtained either by the theory of residues or the theory of orthogonal functions. We have used both methods to determine the Fourier coefficients Q_n and the results agree. The q_n are eigenvalues obtained from the characteristic equation $R(q,a) = 0$, and are purely imaginary since the problem is formulated in terms of modified Bessel functions.

We assume

$$F_n(z,t) = \frac{1}{\pi} \int_0^\infty \frac{W_n(s,z,t)}{s} ds \tag{47}$$

where

$$W_n(s,z,t)=G_n(s,t)\sin(sz)-H_n(st)\cos(sz) \tag{48}$$

On combining these results and substituting the value of $\tau_{r\theta}$ thus obtained in the differential equation

$$\frac{\partial^2\tau_{r\theta}}{\partial r^2} + \frac{1}{\partial} \frac{\partial\tau_{r}}{\partial r^2} - \frac{4\tau_{r\theta}}{r^2} + \frac{\partial^2\tau_{r\theta}}{\partial z^2} - \frac{1}{c_2^2} \frac{\partial^2\tau_{r\theta}}{\partial t^2} \tag{49}$$

we obtain two ordinary linear inhomogeneous differential equations for a $G_n(s,t)$ and $H_n(s,t)$. We solve by Duhamels integral and evaluate $W_n(s,z,t)$. Duhamel's integral will appear inside the integral in Eq. (48). We interchange the order of integration. We obtain on letting $\alpha=t-t_1$,

$$F_n(z,t) = {}_0^t \int_0^\infty \frac{\sin[\alpha\sqrt{s^2-q_n^2}]\sin\beta s ds dt_1}{c_2\, s \sqrt{s^2-q_n^2}} \tag{50}$$

where $\beta = z - T(t_1)$ in the above equation. We differentiate the inner integral partially with respect to β, evaluate the resulting inner integral by means of a known formula,* and integrate with respect to β to regain the original function $F_n(z,t)$. We obtain

$$F_n(z,t) = \tfrac{1}{2}\, Q_n\, c_2\, p_n^2 \int_0^t \int_0^\alpha J_o\, p_n \sqrt{\alpha^2-\beta^2(z,t_1)}\, d\beta dt_1 \tag{51}$$

Where $p_n^2 = - q_n^2$, and is real and positive. Thus we have two quadratures followed by a summation. In practice, the order summation and quadratures should be interchanged to reduce round off error.

*Reference 17, page 472, paragraph 3.876, Eq. (1)

We can readily obtain the response to a step function, then a square wave by superposition and translation. An additional convolution will account for variable torque, which must be obtained from the dynamics of the shell.

VI. DISCUSSION AND CONCLUSIONS

We have suggested methods of improving the accuracy of calculations based on classical analysis without using multiple precision calculations, and which are thus suitable for a group in engineering or applied mechanics. We have obtained formulas for stationary loads, loads moving with constant velocity, and, in the case of torsion, loads moving with arbitrary acceleration. The method presented here for solving the acceleration problem has not been found in the literature and therefore requires careful justification. Additional analysis and considerable programming are required to obtain codes for calculating the stresses and strains. Only then will the results be useful in interpreting strains obtained with instrumented gun tubes. The work is continuing with the time and resources available.

ACKNOWLEDGMENTS

The assistance of Dr. Rurik K. Loder, Ballistic Research Laboratory (BRL), in interpreting Eq. (52) is acknowledged.

J.N. WALBERT

PRESENT ASSIGNMENT: Research Mathematician, Ballistic Research Laboratory.

PAST EXPERIENCE: Mathematician, Material Test Directorate, Aberdeen Proving Ground.

DEGREES HELD: BS University of Delaware, 1970; Master of Science, University of Delaware, 1972; PH.D, University of Delaware, 1975.

K.L. ZIMMERMAN

PRESENT ASSIGNMENT: Mathematician, Ballistic Research Laboratory.

PAST EXPERIENCE: Teacher, Baltimore County, 1965-1966.

DEGREES HELD: BS, Towson State University, 1965.

REFERENCES

1. R. Beeuwkes, Jr., "Stresses in Thick-Walled Cylinders", Watertown Arsenal Laboratory Report 730/419, Watertown Arsenal Laboratory, October 1946. Presented at the Sixth International Congress for Applied Mechanics, Paris, France, September 1946.

2. P.P. Radowski, J. I. Bluhm, O. L. Bowie, Editors, Thick-Walled Cylinder Handbook, WAL Report 893/172, December 1954. See bibliography for additional references.

3. J. Dougall, An Analytical Theory of the Equilibrium of an Isotropic Elastic Rod of Circular Cross Section, Transactions of the Royal Society of Edinburg, Vol. XLIX, Part IV, (No. 17), pages 895-978). (1913).

4. V. K. Prokopov, "Equilibrium of an Elastic Axisymmetrically Loaded Thick-Walled Cylinder," Priladnaya mathematika i mehanika, Vol. XIII, 1949 pages 135-144. Institute of Mechanics of the Academy of Sciences, USSR, FTIO Translation No. J-2589, Aberdeen Proving Ground, Maryland, Translation dated 22 August 1967.

5. A.S. Elder, K. L. Zimmerman, "Stress in a Hollow Cylinder produced by a Step Function of Pressure or Shear", Army Research Office Report 75-1.

6. A.S. Elder, "Formulas for Calculating Bessel Functions of Integral Order and Complex Argument," BRL Report No. 1423, 1968.

7. K.L. Zimmerman, A.S. Elder, A. Depue, "User's Manual for the BRL Subroutine to Calculate Bessel Functions and Integral Order and Complex Argument", BRL Report No. TR-02068, May 1978.

8. A.S. Elder, "Stresses in a Hollow Elastic Cylinder produced by a Step Function of Pressure or Shear," presented at the 16th Polish Solid Mechanics Conference, Krynica, Poland, 26 August - 3 September 1974.

9. A.S. Elder and K.L. Zimmerman, Stresses in a Gun Tube produced by Internal Pressure and Shear, BRL Memorandum Report No. 2495, June, 1975.

10. G.I. Taylor, Strains in a Gun Barrel Near the Driving Band of a Moving Projectile, Report AC-1851, Advisory Council on Scientific Research and Technical Development, 19 March 1942.

11. H. Bucholtz, "The Confluent Hypergeometric Function", translated by H. Lichtblau and K. Wetzel, Springer-Verlag, New York Inc. 1969. See Eq. (10a) page 89 for required infinite integral.

12. E.T. Copson, "An Introduction to the Theory of Functions of a Complex Variable", Oxford University Press, London, 1935. See Ex. 2, page 121, for discussion of zeros of a sequence and its limit function.

13. A.S. Elder, K.L. Zimmerman, J.N. Walbert, "Special Functions for Calculating Elastic Stresses in a Hollow Cylinder," Special BRL Report to be published.

14. S. Timoshenko and J.N. Goodier, "Theory of Elasticity", second edition, McGraw-Hill Book Company, New York, 1951.

15. British Association for the Advancement of Science, Mathematical Tables, Vol. VI, Part I, Bessel Functions of Order Zero and Unity.

16. M. Abramowitz and S.S. Stegun, Editors, Handbook of Mathematical Functions, National Bureau of Standards, Applied Mathematics Series No. 55, June, 1964.

17. I.S. Gradshteyn and I.M. Ryzhik, Tables of Integral, Series, and Products, Fourth Edition, Academic Press, New York, 1980.

18. A.S. Elder, K.L. Zimmerman, E.M. Wineholt, "Ascending and Asymptotic Series for Squares, Products, and Cross-Products of Modified Bessel Functions, Technical Report ARBRL-TR-02391, Ballistic Research Laboratory, February 1982.

TURBULENCE ANALYSIS BY USE OF THE FAST FOURIER TRANSFORM (U)

OSKAR M. ESSENWANGER, Ph.D.
US ARMY MISSILE COMMAND
US ARMY MISSILE LABORATORY, RESEARCH DIRECTORATE
REDSTONE ARSENAL, ALABAMA 35898

1. INTRODUCTION

Standard techniques in the analysis of turbulence include constructing the spectrum of turbulence. In the last three decades most authors have produced the spectrum via the calculation of the autocorrelation function, Tukey (1) or Blackman and Tukey (2). This method disclosed(among other things)a distinct economy in electronic data processing compared with the determination of the coefficients of the Fourier series. In fact, the Fourier transform of the autocorrelation function is the power spectral density, or in short, the spectrum, e.g., Tennekes and Lumley (3), p. 214.

Later Cooley and Tukey (4) introduced the Fast Fourier Transform (FFT) by which spectral values are calculated from the amplitudes of the Fourier series with even less computer time than is used by the autocorrelation method. Thus, "canned" programs of the FFT can be found readily on modern electronic data processing systems while the "old-fashioned" technique via the autocorrelation function is fading away. Some investigators may tacitly accept the postulation that the results from the FFT provide a valid spectrum in the analysis of turbulence. This postulation is examined in the following study.

It will be demonstrated that the FFT can be utilized for turbulence analysis but with certain reservations. The main problem is the calculation of the slope of the spectrum for a longer data series and the scatter of the standardized squared amplitudes for the FFT. The latter may produce outliers which could bias the slope of the spectral density.

It will also be shown that readings of the Gill anemometer (u-v-w). at one-second time intervals produced features in agreement with turbulence theory.

Finally, the FFT implies that the turbulence fluctuations can be superimposed upon the "stationary" profile for simulation studies of Army missile systems.

2. TURBULENCE AND POWER SPECTRUM

The presence of turbulence requires that the energy E(x) in the power spectral density of the one-dimensional velocity (e.g., the wind's u-component) decreases as a function of the (standardized) wave number k:

$$E(k) = \alpha\varepsilon^{2/3}\ k^{-5/3} \qquad [1]$$

This decrease follows the Kolmogorov-Obukhov-Corrsin hypothesis on isotropic turbulence in the inertial subrange; see Tennekes and Lumley (3), p. 266, Priestly (5), p. 61, Hinze (6), p. 194, etc. The Kolmogorov constant α and the dissipation ε (here a constant) are of secondary interest. The reader is referred to the quoted literature for more details on α and ε.

The relationship to the power spectral density L_j is:

$$L_j = \int_{k_1}^{k_2} A_j^2\ (k)\ dk = E(k) \qquad [2]$$

where the wave number $k = j/p$ and p is the basic period. Computation of the power spectral density via the autocorrelation function was introduced by Tukey (1), see also Blackman and Tukey (2) or Panofsky and Brier (7). Power spectrum and Fourier series are associated by:

$$L_j = A_j^2/2\sigma^2 \qquad [3]$$

the squared standardized amplitude of the jth Fourier term; σ^2 is the variance.

In double logarithmic coordinates:

$$\ell n(L_j) = \text{const} - (5/3)\ell n\ k = \text{const} - b\ell n\ k \qquad [4]$$

This is a linear equation with slope $b = -5/3$. Thus, turbulence is different from white noise, a random process, whose spectral characteristic is:

$$\ell n\ L_j = \text{const} \qquad [5]$$

with fluctuations of Lj produced by random errors.

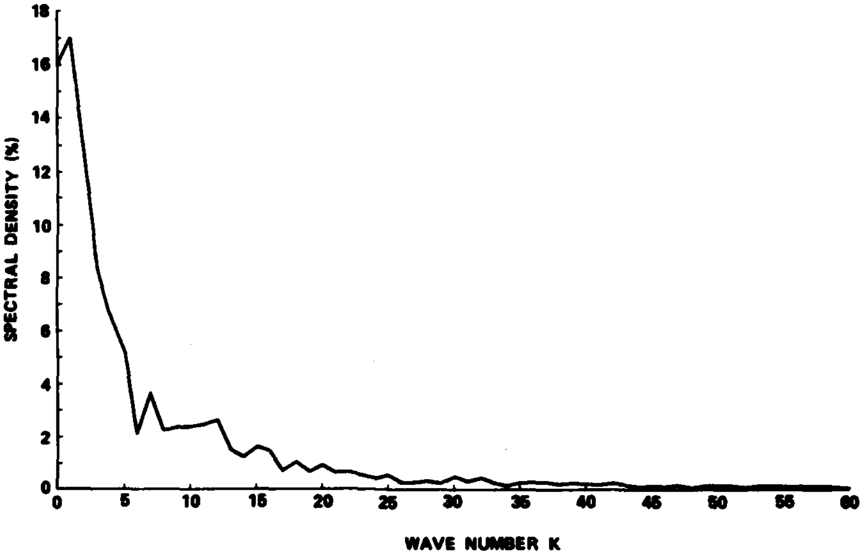


Fig. 1. Power spectrum of the windshear component Δu, 1-second recordings, 19 Aug 74. $\Delta u = u_2 - u_1$, level 2 at 9.1 m, level 1 at 5.5 m, maximum lag 60 seconds.

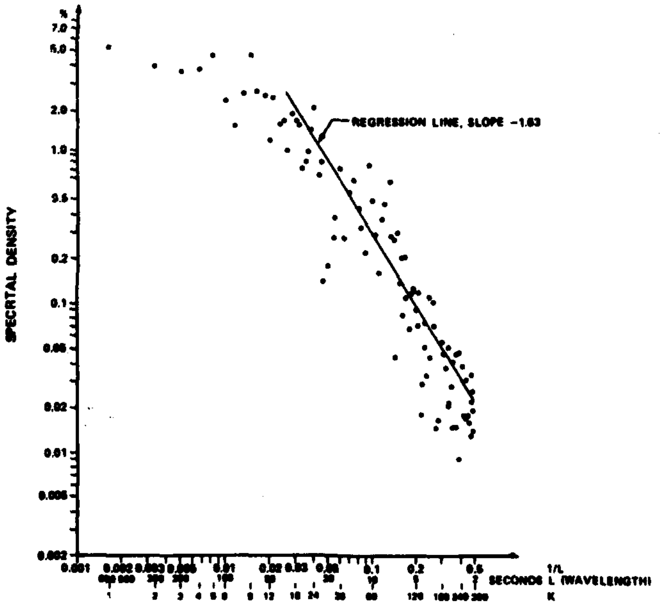


Fig. 2. Power spectrum for data of Figure 1 in double logarithmic coordinates. Maximum lag 300 seconds.

Fig. 1 illustrates the power spectrum for $\Delta u_i = u_{2i} - u_{1i}$ where the u_{ji} are the one dimensional wind components at level j of a meteorological tower. The data have been recorded on 19 August 1974 at one-second time intervals as measured by Gill anemometers (u-v-w), see Gill (8), Drinkrow (9) or Horst (10), on terrain of Redstone Arsenal, see Stewart (11), at 5.5 and 9.1 m height.

As illustrated by Fig. 1, the power spectrum density is not constant which excludes white noise as the generating background. A plot of the spectrum in double logarithmic coordinates (Fig. 2)* with maximum lag m = 300 seconds and calculation of the regression slope reveal that the major part of the spectral density follows almost a linear decline with slope b = -1.63 which is very close to -5/3.

The scatter of the data points is not too large, but deviations from the linear slope are found at low and high wave numbers. The deviations at low wave numbers are well known, e.g., Haugen (12) p.39, p. 169, Hinze (6) p. 203, Tennekes and Lumley (3) p. 270, Nicholls and Reading (13) Peterson (14) and others. Therefore, some authors use filters before turbulence analysis, e.g., Lester (15). The deviation in the region of high wave numbers may be caused by either nonisotropic turbulence such as described by Hinze (6) p. 501 or random noise at the low amplitudes L_j of the spectrum. Although the smoothed spectrum could have been plotted in Figure 2, the author selected to plot the unsmoothed spectral density values to display the scatter.

The computational effort of determining the slope of the regression line and the scatter of points can be reduced for the power spectrum by selecting a smaller maximum lag m. Figure 3 exhibits the spectral density as function of the wave number in double logarithmic coordinates for m = 60. The slope remains at -1.63 but the scatter is smaller than in Figure 2. As expected, the slope is independent of the maximum lag.

3. SPECTRAL DENSITY FROM THE FAST FOURIER TRANSFORM

Cooley and Tukey (4) have introduced the FFT as a technique for the rapid calculation of amplitudes of the Fourier series by electronic data processing. Since the power spectral density can be derived from these amplitudes, the FFT has replaced the Fourier transform of the autocorrelation function in many cases. The user of this substitution must be aware of some differences between these two analytical tools.

The autocorrelation function normalizes the reference of the phase angle of the Fourier terms and includes some smoothing of the data series.

---

* Only a selected number of spectral values have been plotted in Fig. 2 above wave number 60.

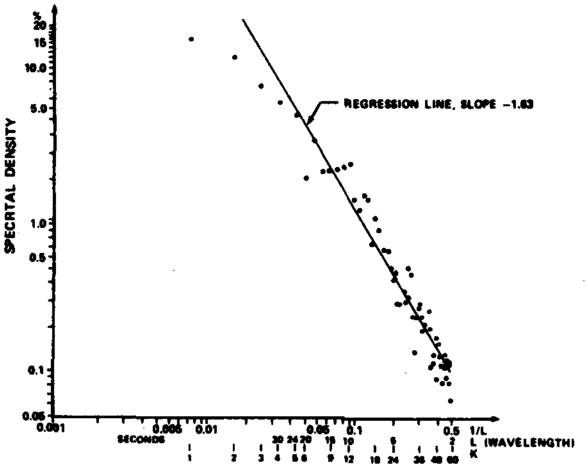


Fig. 3. Power spectrum for data of Figure 1 in double logarithmic coordinates. Maximum lag 60 seconds.

If the data contain "quasi-periodic" waves which appear intermittently, phase angle differences around π (i.e., 180°) between previous and subsequent waves will lead to the diminution or extinction of the amplitude A in contrast to the Fourier transform of the autocorrelation function. This effect may produce some "outliers" in the spectral density of the FFT. Since the FFT is performed on the original data, the spectral amplitudes will also show a larger scatter (see later Figure 6).

The establishment of the power spectrum via the autocorrelation function provides for a choice of the maximum lag m which also determines the basic periods of the analysis. Thus, waves of long length (time cycle) can be lumped together in the wave number k = 0. As previously discussed, the slope is independent of the basic period in the spectrum.

In the FFT the basic period is identical with the length of the data N which also defines the maximum number of terms N/2 or (N-1)/2 whichever is a whole number. Shortening the length of the basic wave can only be accomplished by either truncation of the original data series or by averaging which may suppress waves of small length. Truncation of the autocorrelation function by selection of the maximum lag m is different.

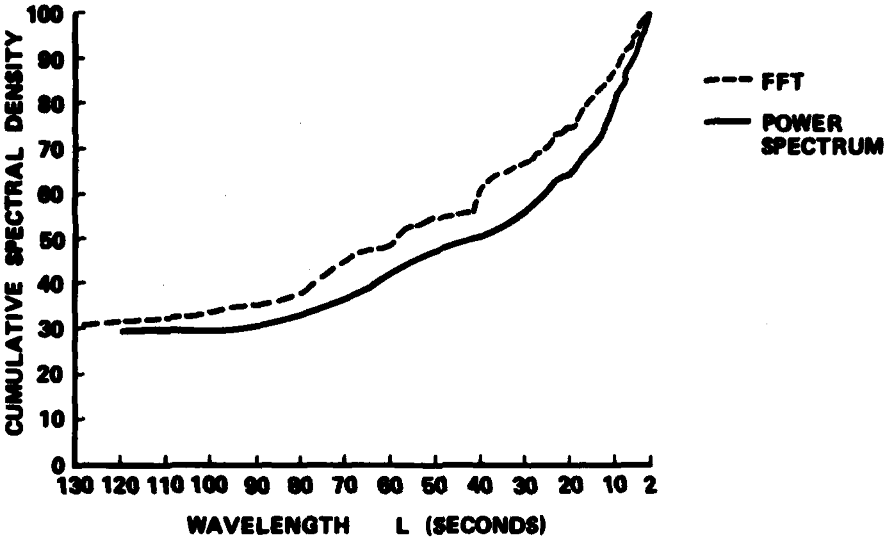


Fig. 4. Cumulative spectral density from autocorrelation (power spectrum) and from FFT for data of Figure 3.

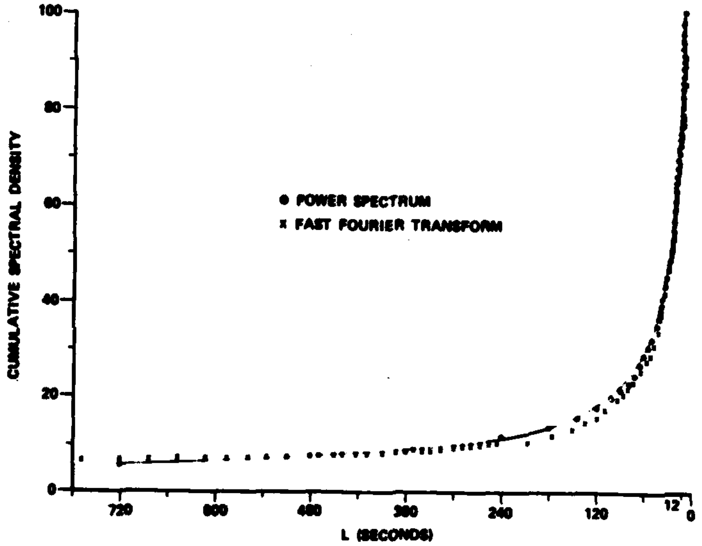


Fig. 5. Cummulative spectral density from autocorrelation (power spectrum) and FFT for data of 25 November 1981.

Consequently, it is no surprise that the cummulative spectral density obtained from the Fourier transform of the autocorrelation function and the FFT for the data of Fig. 2 or 3 exhibit some difference (Fig. 4). A disparity is not always found.

Fig. 5 exhibits the cumulative power spectral density of the Fourier transform spectrum and of the FFT for data recorded on 25 November 1981 at the AFGL Weather Test Facility at Otis AFB, Cape Cod, Massachusetts, measured with a Climatronic Wind Mark I System. As illustrated in this case, the result from the Fourier transform deviates only minimally from the one obtained by the FFT. Although the wave numbers (1-60) in both graphs (Fig. 4 and 5) are the same, the wave length (time cycle) is different because the data have been taken at 6-second intervals. In short, some spectra produced by the two tools may differ and others do not.

Fig. 6 provides the individual (squared and standardized) amplitudes calculated by the FFT for the same data as Fig. 2 and 3. The slope of the regression line b = -1.63 which is identical with the result from the

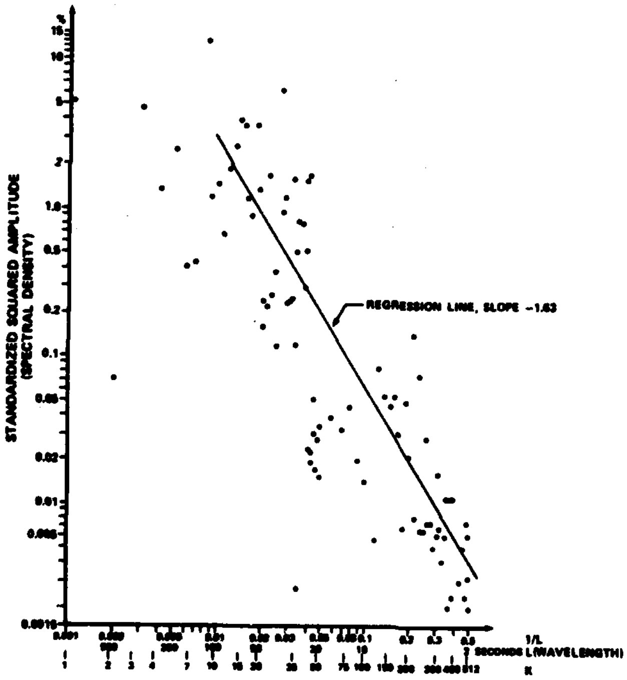


Fig. 6. Spectral density from FFT in double logarithmic coordinates, 19 August 1974 data.

power spectrum but b was not easily obtained. Spectral values of wave numbers with k over 400 and below 10 were truncated, which is equivalent to the application of a bandpass filter.

The weakness of the FFT is the disproportionate weight which high wave numbers have in the ordinary process of calculation of a regression line by least square methods and determination of its slope b. As outlined, the phase angle effect may produce some outliers which also could distort the value of the slope. Two methods are recommended to minimize the effect of disproportionate weight and outliers: truncation or median value regression.

Usually the truncation points (band pass filter) are not known a priori. Thus, they cannot be affixed for electronic data processing before computation of the regression line is made. An iterative truncation process is recommended starting with omitting one or two waves at low wave numbers ($k = 1,2$) and a multiple of waves with high numbers (e.g., for $N/2 = 512$, omit $k=494$ through 512). Then the residual variance and the slope are computed. The residual variance is:

$$v_R = \Sigma(Y_i - y_i)^2/N \qquad [6]$$

where Y_i is the analytical value of $\ell n L_i$ from the regression line and y_i is $\ell n L_i$. The slope b will stabilize after some iterative steps; v_R will decrease.

The iterative process can be combined with a statistical evaluation of the slope b_j from iterative processes j by checking b_j against b_{j+1} or checking v_{R_j} against $v_{R_{j+1}}$. Test methods have been described by Anscombe (16) or Anscombe and Tukey (17).

The fitting of a regression line by ordinary least square methods may not eliminate the biasing effect of outliers, although after some stabilization of b_j and v_{R_j}, excessive deviations $|Y_i - y_i|$ could be omitted and the slope b recalculated. Although a threshold z could be determined and values $|Y_i - y_i| > |z|$ could be excluded, the procedure may be elaborate. A simple but robust method was suggested by Lawson (18) based on the median.

Lawson recommends the division of data into three sections, determining the median of y_i for the extreme sections, and then calculating the regression line from the two median values. This process is simple but may produce an uncertain statistical error. Therefore, for the FFT, this author suggests more than three sections depending on the data length N. An equal number of data in these individual sections is not required.

E.g., sections with a progressive number of data such as k=1-6, 7-18, 19-36, etc. may simultaneously resolve the problem of the excessive weight for waves with a high number of k. The regression line can be fitted to these median values. This reduces the somewhat lengthy process of iteration utilizing the N/2 Fourier coefficients.

It may be of interest to compare v_R for the data presented in Figures 2, 3, and 6 denoted v_{300}, v_{60}, and v_{FFT}, respectively. As expected v_{60} is smallest, $v_{60} = 0.089$. v_{300} is about 20% higher, $v_{300} = 0.108$. $V_{FFT} = 1{,}284$ which is 14 times higher than v_{60} but it is no surprise after examination of Figure 6 and comparing it with Figure 3.

In conclusion, the FFT can be utilized in turbulence analysis with some reservations and precautions. Furthermore, one-second recordings by Gill anemometers disclosed features in agreement with the turbulence hypothesis.

4. THE COMPOSITE WIND PROFILE

The evaluation of the wind effect upon missile systems sometimes requires a detailed wind profile in the microscale for short time intervals over the vertical coordinate. Unfortunately, these microscale observations are available only at special meteorological towers and seldom over 150 m. Thus, data sets must be prepared by analytical methods for simulation studies of missile systems. The results from the analysis presented in the preceding sections, especially in FFT, aid in the construction of these data sets.

It is well known that the wind profile as a function of altitude can be written as:

$$\vec{V}(h) = \vec{V}_s(h) + \vec{V}_t(h) \qquad [7]$$

where V_s represents the "stationary" part and V_t the small scale time and/or space fluctuations of the wind vector. In most cases, only the horizontal components are of importance or interest. Design data of the stationary part have been prepared for profiles of 1, 2 and 10 km by the author (19) at an earlier time. The author together with Billions (20) has also developed a methodology to separate the stationary and non-stationary part from special data measured by Reisig (21). From a power spectrum analysis of the data described by Essenwanger and Billions (20), the author could deduce that the fluctuations of the wind measurements in 15-m height intervals over the vertical range of 1 through 20 km were white noise. In contrast, the present investigations at ground level for a 4-m height difference reveal turbulence behavior of the Δu. Thus, the author postulates from the tentative results that differences of the wind component in the lowest 1 or 2 km may follow turbulence structure rather than white noise, transforming to the latter above that range except in regions with known turbulence such as clear air turbulence. The study continues for confirmation of this postulation. We may now compose the u component for the wind from:

$$u(h) = u_s(h) + \Delta u\ (h) \qquad [8]$$

where $u_s(h)$ denotes the stationary part and:

$$\Delta u(h) = \sum_{j=k_1}^{k_2} A_j \sin(\alpha_{hj} + \beta_j) \qquad [9]$$

Here $\alpha_{hj} = j2\pi\ h/H$ where h is the altitude ($h = 1, \ldots, H$).

In [9] the amplitudes A_j^2 follow the slope $b = -5/3$, i.e.,

$$2\ \ell n(A_k) = B-(5/3)\ \ell nk \qquad [10]$$

B is a constant to be explained below. The phase angles β_j are randomly distributed, i.e., β_j has a rectangular distribution. The waves $k < k_1$ may be considered as part of the stationary profile $u_s(h)$. The upper boundary is $k_2 \leq N/2$ or $k_2 \leq (N-1)/2$.

An equivalent formula is valid for the horizontal (rectangular) v(h) component of the wind.

Formulae [9] and [10] were utilized to simulate the fluctuations of Δu as a function of height (Figure 7). These fluctuations can be superimposed upon $u_s(h)$. This set of data should prove to be better suited for the evaluation of the small scale wind effect than the present technique of assuming a 95% wind profile and superimposing a 95% gust.

The investigations will continue with determination of numerical values for B. Theoretically an initial value can be found for k = 1. Then $B = \ell n(A_1^2)$, but it should be noticed that A_1 is not necessarily identical with the first Fourier term of the FFT as we learn from Figures 2, 3, and 6.

5. TIME AVERAGES AND DIFFERENCES OF WINDSPEED (MICRO-SCALE)

Common windspeed measurements (such as by ordinary cup anemometers) may be considered as time averages because of either a slow instrumental response time or a built-in time integration. In fact, this averaging is intended for conventional measurements to report a "representative" windspeed in synoptic observations. These measurements display a "meso-scale" of turbulence which was presented in the preceding sections. Seldom have attempts been made to relate the two phenomena, but a global evaluation (climatology) of turbulence as required for missile systems is difficult to establish without this important link. This short section serves to offer some solutions from preliminary data.

As previously outlined white noise (a random process for a sequence of independent data) produces a spectrum whose slope in eqn. [4] appears with $b \sim 0$. Persistence leads to $b < 0$. The structure of turbulence is reflected by producing $b = -5/3$. Red noise is another special form of persistence but the slope in the spectrum cannot readily be associated with a fixed numerical value. For smoothed data it may be $b \leq -2$ (Essenwanger and Reiter, 22).

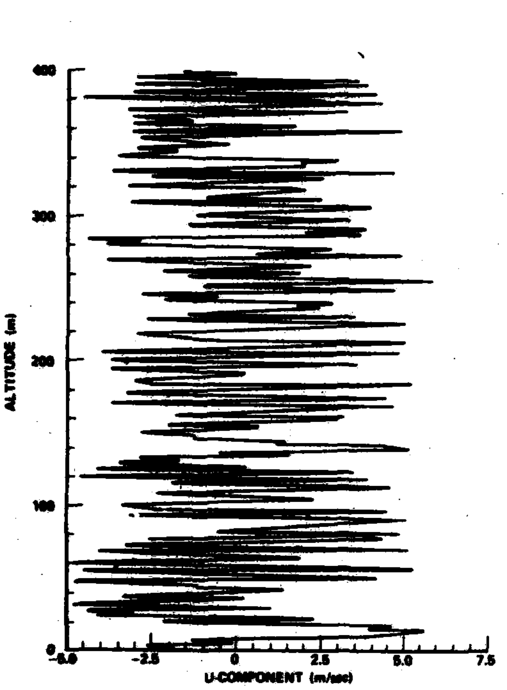


Fig. 7. Simulated data based on Eqns. [9] and [10].

The u and v components of the windspeed data measured on 19 August (see section 2) were subjected to an averaging process of the length 5, 10, 20, and 30 seconds and the spectrum calculated (FFT).

The slope for the u-component at the 5.5 m level appeared as $b = -2.8$ for the 10-second and -3.6 for the 20-second average. These numerical values imply a trend toward red noise.

Examination of the first autocorrelation reveals a drop from 0.97 for 1-second data to between 0.5 to 0.6 for the 30-second averages which indicates a looser connection between individual data of the sets. It is the structure of the autocorrelation, however, and not the first lag correlation which determines the link to red noise. As an example the sequence of the autocorrelation coefficients from lag 0 to lag 10 are exhibited for 1-second observations and 10-second averages (Table 1). Red noise requires:

$$r_i = (r_1)^i \qquad [11]$$

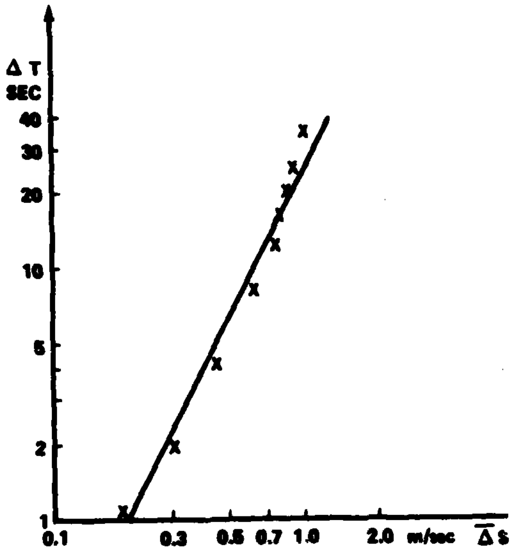


Fig. 8. Average windspeed $\overline{\Delta S}$ as function of time interval.

Table 1. Autocorrelation Function of u-component and Red Noise Autocorrelation Series

Lag	1-Sec.	Red N.	20-Sec.	Red N.
0	1.0	1.0	1.0	1.0
1	.974	0.974	.679	.679
2	.936	0.949	.459	.461
3	.900	0.924	.383	.313
4	.866	0.900	.272	.212
5	.836	0.877	.151	.144
6	.809	0.854	.103	.098
7	.785	0.832	.081	.065
8	.762	0.810	-.026	.045
9	.738	0.789	.028	.031
10	.715	0.768	.102	.021

which is identical with a first order Markov chain (Essenwanger, 23). We notice that the autocorrelation coefficients of 1-second data (u-component, level 5.5 m, 19 Aug. 74) drop more quickly than expected from red noise. In contrast, the autocorrelation coefficients for the 20-second averages are in line with expectations from red noise. The exceedance at lags 3 and 4 and the outlier at lag 8 seem to be associated with the presence of a quasi-cycle of around 2½ minutes during the 40-minutes when the data sample was taken on 19 Aug. 74.

Table 1 confirms the assumption of a red noise pattern in time averaged windspeeds. Thus, power spectra of conventional wind measurements could be interpreted as red noise. Reference to 1-second turbulence structure can then be achieved by determining the constants of eqn. [10] from the mid range of these spectra from "meso-scale" data rather than from fitting the first amplitudes (low k) as suggested for micro-scale turbulence data.

A second behavioral fact must be considered. Figure 8 displays the average difference of the windspeed $\overline{\Delta S}$ as a function of the time intervals ΔT, where:

$$\Delta S = [(\Delta u)^2 + (\Delta v)^2]^{\frac{1}{2}} \qquad [12]$$

Figure 8 discloses a linear relationship in the time range from

1 to 40 seconds differences with a slope of 0.5. Since the first 30 seconds were of primary interest here, the continuation of the data after the 40-second time interval is not shown, but it stabilizes at a constant value for longer time intervals. This result in micro-scale resembles earlier work by Neumann (24) who has demonstrated with customary wind data at the "meso-scale" for readings half an hour apart that the relative variability tends to a constant value as the time interval increases.

If the time structure as disclosed by Fig. 8 is not automatically met during the derivation of analytical data by eqn. [10],it may be necessary to find a solution to include the time structure on a vertical scale. This investigation is continuing.

The results for the 19 Aug 74 data sample were not completely conclusive, and further studies on data samples such as the 25 Nov 81 data will continue. However, one tentative result can be deduced. The numerical value of the slope in Figure 8 is 0.5. This is the same numerical value as derived for the windshear relationship for small intervals (Essenwanger and Reiter, 22,and Essenwanger, 25). Essenwanger and Reiter (22) could interpret the slope of 0.5 as a mixture between smoothed data and turbulence fluctuations. If this interpretation can be applied to the data of Fig. 8, a separation of the "stationary"and"non-stationary" part such as required in eqn. [7] is already the solution, and the time interval relation would implicitly appea.

6. CONCLUSION

The author has shown that the Fast Fourier Transform is a useful tool in turbulence analysis and the lack of "canned" programs of the Fourier transform of the autocorrelation function poses no serious problem. The user is cautioned, however, to be aware of the differences between the spectra produced by these two techniques.

A data sample procured with the Gill anemometer recorded at one-second time intervals reflect agreement with the turbulence hypothesis (Fig. 2,3).

The FFT has the advantage that turbulence data can be readily expressed as a Fourier series. The amplitude relationship is expressed by eqns. [4] and [10] while the phase angles are randomly distributed. Although a set of phase angles can be constructed from random generators in electronic data processing, a simple technique would be a substitution from a set of empirical turbulence data.

The author has given an example of an analytically produced "non-stationary" set of data which could be superimposed on the "stationary" wind profile. These inferred data sets can be established where special tower measurements are not available such as for altitudes beyond tower measurements and/or for geographic locations without tower measurements. These composed sets of wind data are better suited than present techniques for the assessment of the turbulence impact upon effectiveness, instrumental or missile sensitivity, and field use of systems such as DAFFR, Assault Breaker, CSWS, etc. by combining the "stationary" and "nonstationary" effect into one data set.

ACKNOWLEDGMENT

The author expresses his appreciation to Dr. Dorathy A. Stewart for her critical comments and the editorial review of the manuscript. Mrs. Alexa M. Mims deserves most of the credit for the preparation of the computer programs. My thanks go to Mr. Larry J. Levitt for his assistance in the preparation of the figures. The simulated data (Fig. 7) were prepared by Dr. Shi Tsan Wu. Finally, Mrs. Louise H. Cooksey deserves the credit for her patience and expeditious typing of the manuscript, and to CPT James C. Weyman, Air Force Geoph. Lab., for the special data of Fig.5.

REFERENCES

1. Tukey, J. W., 1949. The Sampling Theory of Power Spectrum Estimates. NAVEXOS-P-735, Woods Hole. Office of Naval Research, p. 47-67.

2. Blackman, R. B. and J. W. Tukey, 1958. The Measurement of Power Spectra. Dover Publ. Inc., New York, pp 190.

3. Tennekes, H., and J. L. Lumley, 1973. A First Course in Turbulence. MIT Press, Cambridge, Mass. pp 300.

4. Cooley and J. W. Tukey, 1965. An Algorithm for the Machine Calculation of Complex Fourier Series. Math. Comput. Vol. 19, p. 297-301.

5. Priestley, C. H. B., 1959. Turbulent Transfer in the Lower Atmosphere. University of Chicago Press, Chicago, pp 130.

6. Hinze, J. O., 1959. Turbulence, An Introduction to its Mechanism and Theory, McGraw Hill, New York, pp 586.

7. Panofsky, H. A. and G. W. Brier, 1958. Some Applications of Statistics to Meteorology. Pennsylvania State University, pp 224.

8. Gill, G. C., 1975. Development and Use of the Gill UVW Anemometer Boundary-Layer Meteorol., Vol. 8, p. 475-495.

9. Drinkrow, R., 1972. A Solution to the Paired Gill Anemometer Response Function. J. Appl. Meteorol., Vol. 11, p. 76-80.

10. Horst, T. W., 1973. Corrections for Response Errors in a Three-Component Propeller Anemometer. J. Appl. Meteorol., Vol. 12, p. 716-725.

11. Stewart, D. A., 1975. Turbulence Measurements from the Army Gas Dynamic Laser Range. US Army Missile Command Report RR-75-8, pp 25.

12. Haugen, D. A., 1973, editor. Workshop on Micrometeorology. Am. Meteor. Soc., Boston, pp 392.

13. Nicholls, S. and C. J. Readings, 1981. Spectral Characteristics of Surface Layer Turbulence Over the Sea. Qu. J. Roy. Meteorol. Soc. Vol. 107, p. 591-614.

14. Peterson, E. L., 1976. A Model for the Simulation of Atmospheric Turbulence. J. Appl. Meteorol. Vol. 15, p. 571-587.

15. Lester, P., 1972. An Energy Budget for Intermittent Turbulence in the Free Atmosphere. J. Appl. Meteorol., Vol. 11, p. 90-98.

16. Anscombe, F. J., 1967. Topics in the Investigation of Linear Relations Fitted by the Method of Least Squares. J. Roy. Stat. Soc., B, Vol. 29, p. 1-52.

17. Anscombe, F. J. and J. W. Tukey, 1963. The Examination and Analysis of Residuals. Technometrics,Vol. 5, p. 141-160.

18. Lawson, J. S., 1982. Application of Robust Regression in Designed Industrial Experiments. J. Qual. Techn., Vol. 14, p. 19-33.

19. Essenwanger, O. M., 1973. Wind Profile Parameters in the Ground Layer and Troposph. at Mid Latitudes. US Army Missile Comm.RR-73-6, pp 27.

20. Essenwanger, O. M. and N. S. Billions, 1965. The Stationary and Non-Stationary Wind Profile. Pure and Appl. Geoph. Vol. 60, p. 160-166.

21. Reisig, G. H. R., 1956. Instantaneous and Continuous Wind Measurements up to the Higher Stratosphere. J. Meteorol. Vol. 13,p 448-455.

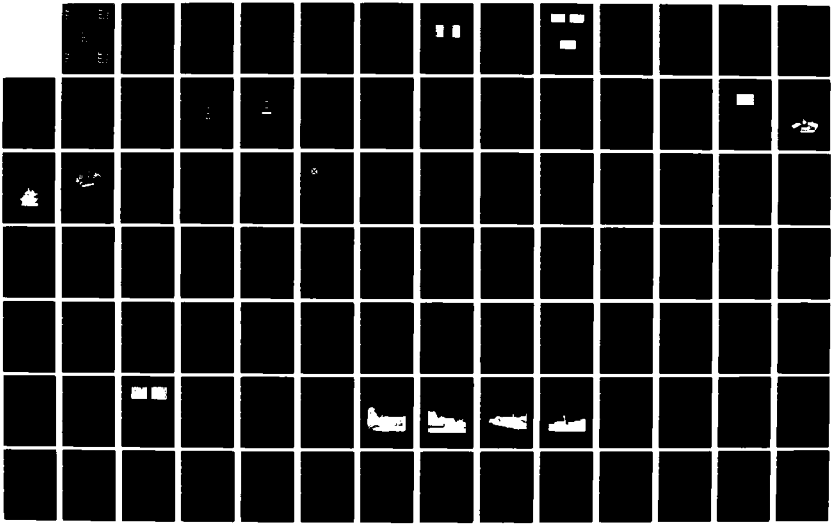
22. Essenwanger, O. M., and E. R. Reiter, 1969. Power Spectrum, Structure Function, Vertical Wind Shear, and Turbulence in Troposph. and Stratosph. Arch. Meteorol. Geoph. Bioklim, A, Vol. 18, p. 17-24.

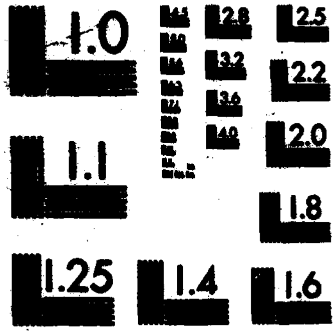
23. Essenwanger, O. M., 1980. On Red Noise and Quasi-Periodicity in the Time Series of Atmospheric Temperature. Statistical Climatology, p. 165-181. Editors: S. Ikeda, E. Suzuki, E. Uchida and M. M. Yoshino. Elsevier, Amsterdam, pp 388.

24. Neumann, J. 1967. Wind Variability for Time Intervals of Medium Length. J. Appl. Meteor., Vol. 6, p. 587-589.

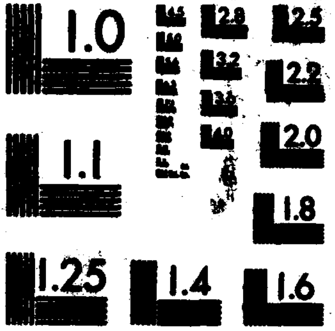
25. Essenwanger, O. M., 1963. On the Derivation of Frequ. Distrib. of Vector Shear Values for Small Shear Intervals. Geofis. Pura Appl. Vol 56, p. 216-224.

AD-A120 811 PROCEEDINGS OF THE 1982 ARMY SCIENCE CONFERENCE HELD AT 5/6
THE UNITED STATES..(U) DEPUTY CHIEF OF STAFF FOR
RESEARCH DEVELOPMENT AND ACQUISITIO.. 18 JUN 82
UNCLASSIFIED F/G 5/2 NL

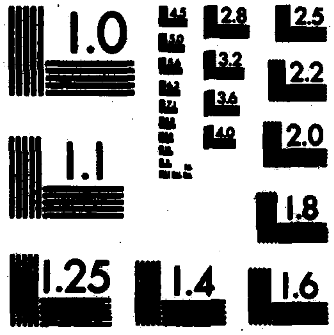



1.0
4.5
5.0
5.6
6.3
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

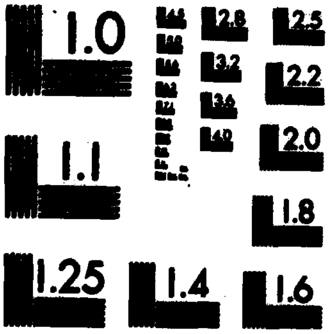
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

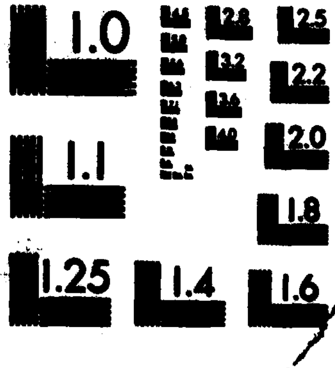
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


1.0
4.5
5.0
5.6
6.3
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

MECHANISMS OF IMPROVED BALLISTIC FABRIC PERFORMANCE (U)

F. FIGUCIA *
C. WILLIAMS
B. KIRKWOOD
W. KOZA

U.S. ARMY NATICK RESEARCH AND DEVELOPMENT LABORATORIES
NATICK, MASSACHUSETTS

Protection of U.S. infantry personnel in ground warfare environments requires constant technical vigilance on the part of Government scientists to assure that the most advanced state-of-the art is being implemented for the benefit of the soldier. This paper addresses that requirement as related to protection of personnel against fragmenting munitions.

Though the level of ballistic protection already achieved and currently being fielded exceeds anything ever developed in the past, the potential exists for even greater improvement in protection against fragmenting threats, and the need exists for protection against new threats. New developments in weapons technology utilizing surface coatings, novel shape factors and higher density materials continually challenge the level of protection achievable with existing armor systems.

A more intensified scientific approach must be devoted to the ballistic problem to provide adequate protection against the rapidly changing threat patterns in the field. The problem is a difficult one for a number of reasons. Fibrous materials have proven their worth both ballistically and functionally, but relatively little scientific information has been developed in this area as compared to other material disciplines. Penetration mechanics of such materials are physically unique and somewhat confounding, and are not clearly defined. Finally, scientific contributions from the private sector have been essentially non-existent due to the marginal commercial attraction of armor items.

It is expected that through this paper a firmer scientific base will be established and that the future use of this technology will eventually lead to a better understanding of the physical interaction between flexible fabric and penetrating projectile.

GENERAL APPROACH

The approach taken in this study is to analyze the projectile impact of a fabric in logical steps, and perform laboratory experiments to quantify the various physical responses which are active. The event involves a number of fiber characteristics and penetration mechanisms interacting in microseconds of time.

The process begins with the initial contact of a projectile upon an orthogonally woven fabric. This sets up a potential exchange of projectile kinetic energy into yarn tensile stress-strain energy. As long as the tensile mode of response remains in effect, the energy of the projectile is translated longitudinally along the yarns in continuous pulses which can be thought of as wavelets of stress-strain energy. These travel at velocities which vary with the crystalline character of the target material. The amount of yarn which becomes involved through this mechanism increases from the time of initial contact until the projectile penetrates the fabric and the strain is released.

At the same time as the in-plane exchange of energy takes place, projectile energy is also dissipated through fabric transverse motion. The forward thrust of the projectile carries target fabric out of its original plane. The displaced material takes the shape of a pyramid (1,2) whose base and height dimensions continually increase until penetration occurs.

The problem is approached through an analysis of the energy dissipation in these two modes (planar and transverse). Theoretical energy absorption is derived after quantification of appropriate parameters. The computation for planar energy encompasses penetration time, wave velocity, number of contacted yarns, and yarn energy absorption potential. Computation of the transverse component of energy involves the shape and dimensions of the fabric distortion, the average missile/fabric velocity, and the mass of the fabric displaced.

These computed energy absorptions are then compared with that measured in actual tests. The agreement obtained will establish the importance of the studied parameters and their application to future efforts in the field.

MATERIALS

The main thrust of this analysis was toward the behavior of woven fabrics as opposed to other constructional forms.

Fabrics representative of two polymer types were studied; nylon 66 (polyamide) and Kevlar 29 (polyaramid). Fabrics were selected with constructional features as nearly identical as possible. Characteristics of the two materials are tabulated below.

Table 1

Characteristics of Kevlar 29 and Nylon 66 Fabrics Used for Analysis

	Kevlar 29	Nylon 66
Polymer Type	Polyaramid	Polyamide
Yarn Fineness, Tex (Denier)	111 (1000)	117 (1050)
Weave	2 x 2 Basket	2 x 2 Basket
Weight, g/m^2 (oz/yd^2)	431 (12.7)	505 (14.9)
Yarns/cm, w x f	19.7 x 17.3	20.5 x 17.3
Yarn Crimp (%), w x f	10.1 x 1.0	12.5 x 4.5

Experiments were performed to obtain a) actual measurement of fabric energy absorption under ballistic impact and b) data with which to compute this energy based on hypothetical energy transfer modes. Measurement of actual energy absorption is described under Fabric Energy Absorption. The remaining experiments for yarn stress-strain properties, strain wave velocity, penetration time and transverse displacement were for computed data.

Fabric Energy Absorption

Tests to measure the energy absorption of single fabric layers were performed in two series using the two different fabric clamping arrangements depicted in Figure 1. The first uses a pair of aluminum plates between which the fabric is clamped. The target area is a 22.9 cm

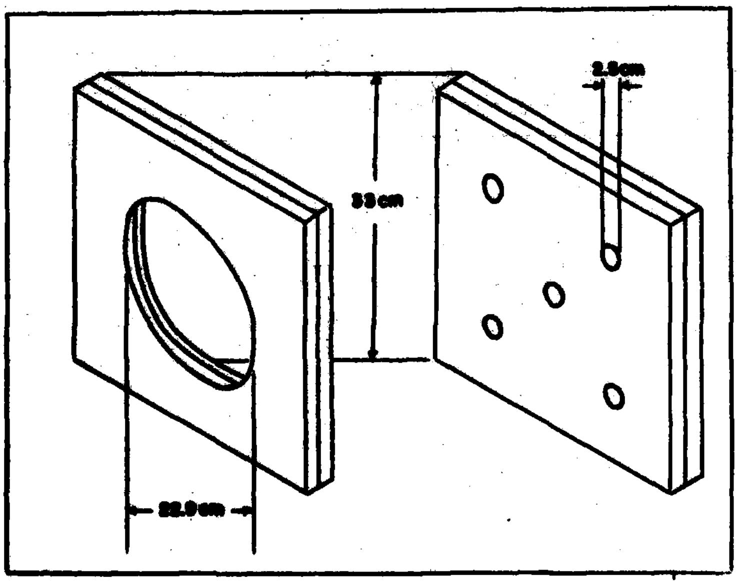


Figure 1. Fabric Clamps Used For Energy Absorption Tests.

diameter circle which has been cut out of the plates. The second clamp system has the same outside dimensions but uses thicker, heavier aluminum plates. Rather than using a single target circle, it has five 2.5 cm circles cut from the plates. This clamp was discussed in more detail in a previous report (3).

Energy absorption for Kevlar and nylon fabrics was measured using both clamping arrangements at projectile velocities ranging from stopping velocity to about 500 m/s. The test apparatus and methodology were described in a previous study (4). Briefly, the energy absorbed using either fabric clamp is computed from accurate measurement of projectile velocity before and after penetration using electronic sensing devices and digital counters. These velocities are converted to kinetic energy, the difference ($E_{in} - E_{out}$) being the projectile energy absorbed or transmitted to the fabric.

Yarn Stress-Strain

Tensile stress-strain curves were generated for individual yarns using an Instron Tensile Tester. A gage length of 12.7 cm was used with a crosshead speed of 5 cm/min resulting in a straining rate close to 40%/min.

Data at higher rates of strain will be referenced in the Discussion portion of this paper. These data were developed prior to this study using pneumatically driven tensile test equipment to achieve rates in the order of 3×10^5%/min. The best estimates of externally applied straining rates under ballistic impact are in the range of 10^6 to 10^7%/min.

Strain Wave Velocity

Measurement of strain wave velocity in the fabrics was made using an extension of a method developed by Koza (5) for similar measurements in single yarns. Fabric target samples were arranged as shown in Figure 2.

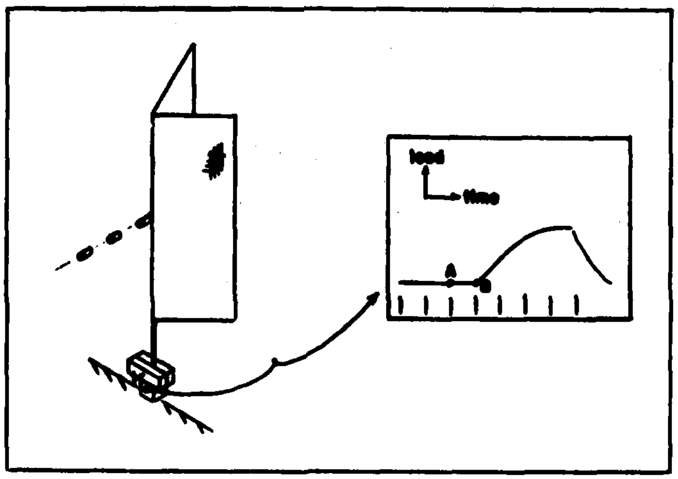


Figure 2. Test Setup For Fabric Strain Wave Velocity.

The fabric was clamped at the top but only 10 yarns extending from the bottom edge were clamped below. These were the ends of the same yarns to be contacted by the projectile at the center of the fabric. A piezoelectric crystal load cell was fastened to the bottom clamp as shown and its output directed to an oscilloscope. Knowing the distance between the load cell and the point of contact of the projectile with the fabric, it was only necessary to measure the time of travel of the wave over this distance to obtain wave velocity. To accomplish this the test was performed with the room in darkness and with a multiflash generator and camera set up to photograph the flight of the projectile. The first flash was synchronized with the sweep of the dual oscilloscope beams. The upper beam recorded load cell response vertically vs time horizontally. The lower beam gave a spike signal for each flash. The example in Figure 2 shows that the projectile contacted the fabric on the 3rd flash which on the load-time recording is a point of no apparent load (point A). The response of the load cell occurs at point B a measurable time later. The distance AB represents the travel time of the wave.

Wave velocities were measured in this way for both nylon and Kevlar fabrics in both warp and filling directions. Projectile velocities of approximately 305 m/s were used for these tests.

Penetration Time

The time elapsed during penetration was measured over the same range of projectile velocities as used for the fabric energy absorption tests. The fabric clamp with the 22.9 cm diameter hole in the center was modified to facilitate this experiment. A 0.6 cm slot was bored upward from the bottom edge of the clamp on the inside face of each plate. This relieved about 10 yarns from the clamping pressure during testing. These were the same yarns contacted by the projectile at the center of the target. They were extended downward to the clamp and load cell which were coupled together below the target as shown in Figure 3. Penetration time (t) was

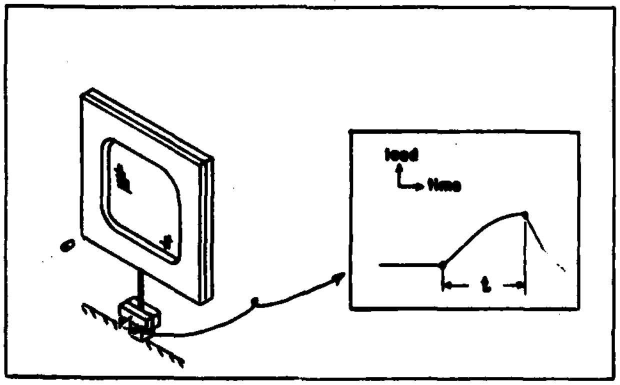


Figure 3. Setup For Penetration Time Tests.

measured from the oscilloscope force-time trace as that time from the initial response of the load cell to maximum load (penetration). Test specimens were oriented with the filling yarns vertical. Since these had less crimp than the warp yarns, sharper loading curves were produced.

Tranverse Displacement

The component of projectile energy transmitted transversely in the fabric is a function of the size of the fabric pyramid formed during penetration. The pyramid dimensions were measured from photographs of the exit side of the fabrics taken at projectile penetration (Figure 4).

Nylon

Kevlar

Figure 4. Transverse Distortion of Nylon and Kevlar Fabrics at Penetration

Measurement of the base dimensions of the pyramid were made directly from the photographs which were actual size. The pyramid heights were measured from the peak of the extended pyramid back to the location of the point of initial contact of the projectile. This point was established as the intersection of the diagonals drawn from the corners of the pyramid base.

Projectile velocities for these tests were at the point at which penetration just takes place and the largest pyramids occur.

DISCUSSION

The energy absorbed by single layer fabrics under transverse impact is propagated by two mechanisms: a) tensile stress-strain work performed by the individual yarns acting within the fabric plane, and b) transverse kinetic motion of the displaced fabric mass. The in-plane energy absorption, E_p can be computed as the product:

$$E_p = 2(t)(c)(n)(E_y) \qquad \text{(Equation 1)}$$

where: E_p = Fabric in-plane energy absorption (J)
t = Penetration time (s)
c = Strain wave velocity in fabric (m/s)
n = Number of yarns under strain
E_y = Yarn energy potential (J/m)

The constant "2" accounts for the propagation of energy in both directions from the point of impact along the contacted yarns.

The transverse energy component, E_T, can be estimated from pyramid dimensions and projectile/fabric velocities as:

$$E_T = \frac{A\,\underline{W}\,\bar{v}^2}{2} \qquad \text{(Equation 2)}$$

where: E_T = Transverse energy (J)
A = Area of pyramid base (m^2)
$\underline{W}$ = Fabric weight (kg/m^2)
$\bar{v}$ = Average velocity of projectile during penetration (m/s)

The input parameters to the above equations can vary substantially depending upon fabric construction and ballistic test conditions. These variables along with other relevent factors will be discussed in the following sections.

Failure Modes

The prediction of in-plane energy absorption by equation 1 is applicable as long as the resistance to penetration is through tensile work done by the yarns. It is probably fair to assume that tensile straining of yarns takes place, to some extent at least, in ballistic impacts of all fibrous materials. However the impinging energy will ultimately cause failure through the path of weakest resistance of a given material. Two such modes have been observed in previous ballistic experiments. They are failure due to low heat resistance (melting failure) and failure due to excess stiffness (brittle failure). These are illustrated in the scanning electron micrographs of Figure 5.

The polyamide fibers are representative of the nylon 66 used in this study. Melting is obvious at the broken fiber ends. Melting has also been noted on the surface of unbroken fibers outside the perimeter of missile contact.

Brittle failure is illustrated by an experimental polyaramid which exhibited poor ballistic resistance despite its high strength. The sharp transverse failure shown is typical of other high strength brittle fibers such as glass and steel.

Polyamide

Experimental Polyaramid

Figure 5. Fracture Morphology of Fibers with Poor Heat and Shear Resistance.

Kevlar, which is also a polyaramid with brittle characteristics, apparently maintains sufficient shear resistance to withstand the transverse pressure of the missile. Figure 6 shows Kevlar fibers after ballistic impact. Neither of the failure modes of Figure 5 are evident. The fibrillated broken ends are typical of those from laboratory tensile tests.

Figure 6. Fibrillation of Kevlar Fibers After Ballistic Impact.

Yarn Stress-Strain Energy

The nylon and Kevlar yarns used in this study exhibit the characteristic stress-strain behavior shown in Figure 7 when tested at a quasi-static straining rate of 40%/min. Previous studies (6, 7) have shown that for tensile straining rates of 3 x 10^5%/min differences exist in the ultimate strength, ultimate extension and initial modulus properties.

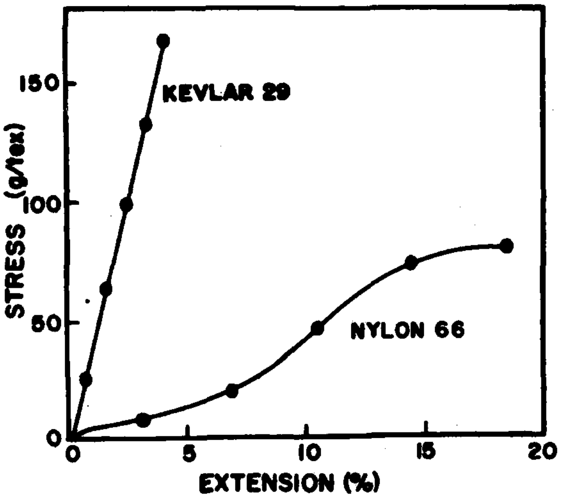


Figure 7. Stress-strain Behavior of Nylon and Kevlar Yarns.

However the energy absorbed by the yarns, as measured by the areas under the stress-strain curves, remains essentially the same as at the slow strain rate. Similar trends in tensile behavior have also been noted for a wide variety of other polymeric materials tested at high and low rates of tensile strain (8). Mechanical properties of nylon and Kevlar yarns at the two strain rates are shown in Table 2.

It is assumed for purposes of this study that yarn energy absorption at ballistic strain rates (estimated at 10^6-10^7%/min) remains at the levels of Table 2. The conversion of Table 2 energy data to the units of equation 1 give E_y values of 3.9 J/m for Kevlar and 7.9 J/m for nylon.

Table 2

Tensile Properties of Nylon 66 and Kevlar 29 Yarns at Low and High Rates of Strain

	Nylon 66		Kevlar 29	
	L*	H*	L	H
Break Stress (g/Tex)	79	90	169	198
Break Extension (%)	19	16	4.3	3.7
Initial Modulus (g/Tex)	387	540	3700	5400
Break Energy (g/Tex)	7.4	7.3	3.5	3.7

*L = 40%/min strain rate
*H = 3 x 10^5%/min strain rate

Strain Wave Velocity

The longitudinal propagation of stress along single yarns has been treated theoretically by many researchers (9, 10, 11). In general, strain waves travel in continuous pulses from the point of stress (contact of projectile) outward along the yarn with a velocity (c) which is related to yarn modulus (E) by $c = (E/k)^{\frac{1}{2}}$ when E is expressed in textile linear density units. They continue until they lose their intensity or they reach a boundary such as a test clamp from which they are reflected back toward the point of impact. Some feel that this causes a multiplication of stress at the point of reflection which, given sufficient time, will cause failure at these locations. This may be a contributing factor to the change in response from tensile to other possible modes of failure discussed earlier.

Application of yarn wave theory to the fabric case is complicated by such reflections from successive yarn crossovers and also by the crimp formed in the yarns during weaving. Therefore it was decided that direct measurement of the wave velocity in the constructed fabric was the only appropriate method of analysis for this study. Tests were performed in both warp and filling fabric directions. Results were compared to previously obtained data for single uncrimped yarns and are shown in Table 3.

Table 3. Strain Wave Velocity Measured in Fabric and Free Yarn

	Strain Wave Velocity (m/s)	
	Kevlar 29	Nylon 66
Warp	900	900
Filling	5000	1000
Free yarn	7600	2400

Velocities are greatest in the uncrimped free yarns for both materials. Once these same yarns are woven into fabric however, the velocities are reduced due to fabric structural effects. The extent of this structural interference on wave velocity varies with fabric direction. It can be related to crimp in the manner shown in Figure 8. The superior wave propagation characteristic of uncrimped Kevlar is gradually reduced until this advantage is totally lost at approximately 10% fabric crimp. This result emphasizes the importance of fabric structural effects on ballistic performance. It has been shown(3,4) that Kevlar fabrics woven with a minimum of interlacings and yarn crimp (satin weave) perform significantly better than tighter weaves. This constructional advantage which had previously been attributed to improved transverse mobility is now shown to be related to in-plane energy parameters.

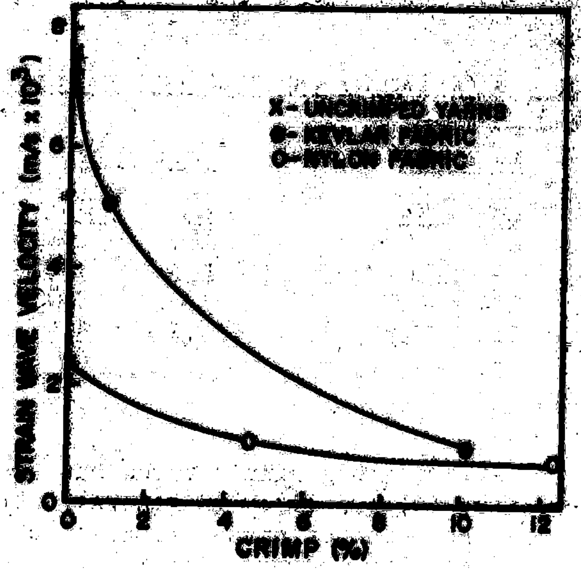


Figure 8. Effect of Fabric Crimp on Strain Wave Velocity.

Penetration Time

Penetration time defines the period during which the yarns are under stress and therefore the distance travelled by the strain waves. Penetration times (t) were measured for input to equation 1 over the series of projectile velocities shown in Figure 9. Kevlar shows a distinct dependency on projectile velocity which reflects the reduction in extension with increased strain rate observed in Table 2.

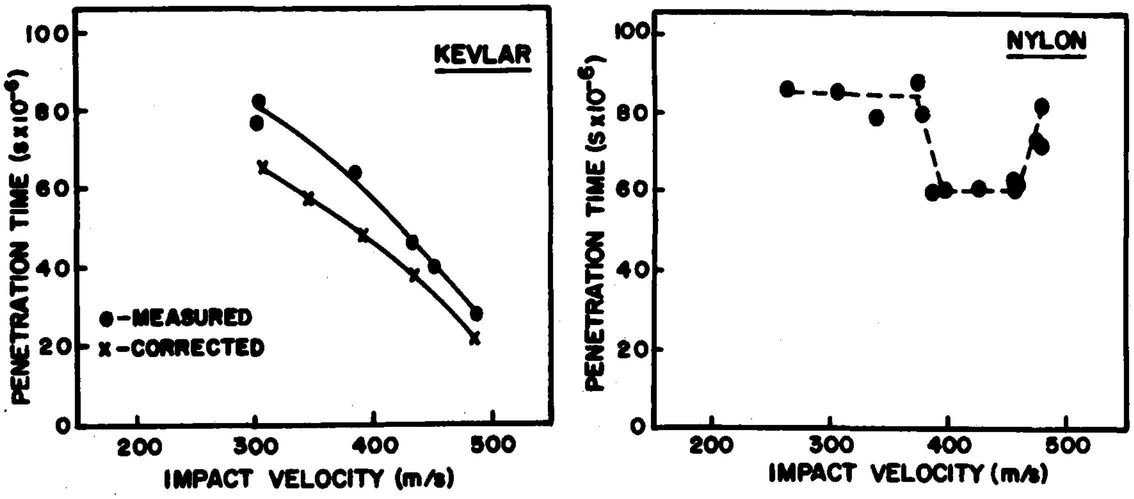


Figure 9. Effect of Projectile Velocity on Time of Penetration for Kevlar and Nylon Fabrics.

Nylon does not show the rate dependency of Kevlar but rather an erratic pattern which may have some explanation based on heat phenomena. No further attempt is made to explain these results at this time.

Alteration of the clamp faces to facilitate measurement of penetration time eliminated the barrier effect of the clamp on the waves traveling in one of the four directions away from the impact point. This undoubtedly increased the penetration time over that which would apply for a totally clamped fabric. A correction of 25% was used to reduce the values for Kevlar. A similar correction for nylon was not considered necessary since it does not appear that the tensile mode was active.

Transverse Energy

The amount of projectile energy lost through transverse fabric motion E_T is estimated from equation 2, which measures the kinetic energy of the displaced fabric mass at the average penetration velocity.

Measurements were made of the fabric pyramids formed during impact at velocities near that at which penetration just occurs and the largest pyramids are formed. E_T values obtained under these conditions were used to estimate those at higher velocities through the use of data of Roylance (2). This resulted in a reduction of transverse energy absorption with increasing impact velocity (Figure 10).

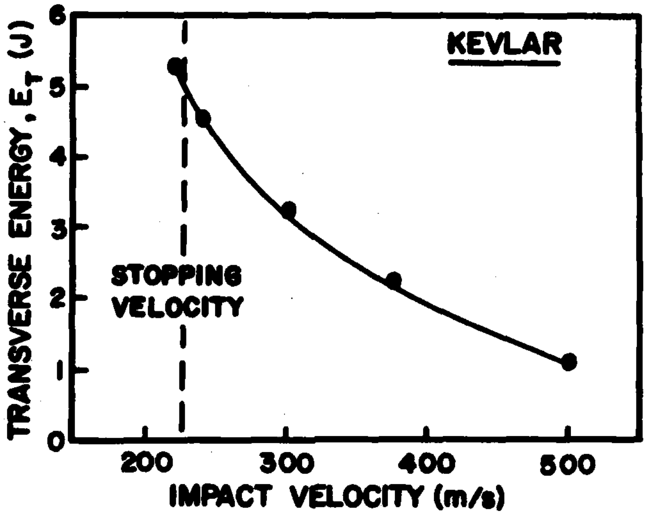


Figure 10. Effect of Projectile Velocity on Transverse Energy Absorption of Kevlar.

Theoretical vs. Experimental Energy Absorption

Single layer fabric samples were tested for energy absorption over a range of projectile velocities using 2.5 and 22.9 cm diameter clamps. Figure 11 shows the effects of the clamp boundaries on the results. The maximum energy absorbed by the two materials at the velocity of initial penetration (left most point) is reduced about 50% in going from the larger to the smaller target. This difference becomes smaller however as projectile velocity increases due to the accompanying reduction in penetration time and wave propagation distance. The reduction in energy absorption (50%) is not proportional to the ratio of target diameters (11%). Considerable yarn strain was evident in all directions beyond the 2.5 cm boundaries, indicating that the strain wave intensity was sufficient to force through the nearby boundaries resulting in more yarn involvement and energy absorption levels greater than predicted.

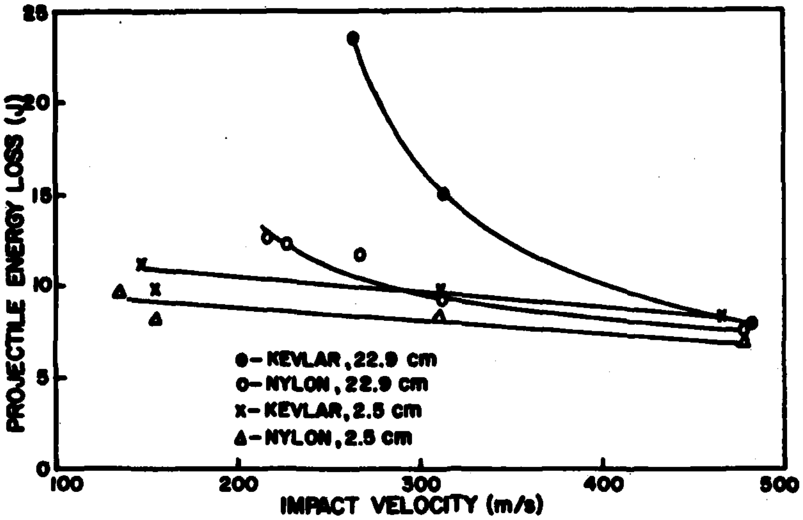


Figure 11. Effect of Clamp Boundaries on Experimental Energy Absorption.

The theoretical energy trend for the 22.9 cm Kevlar targets showed close agreement with that obtained experimentally (Figure 12). This

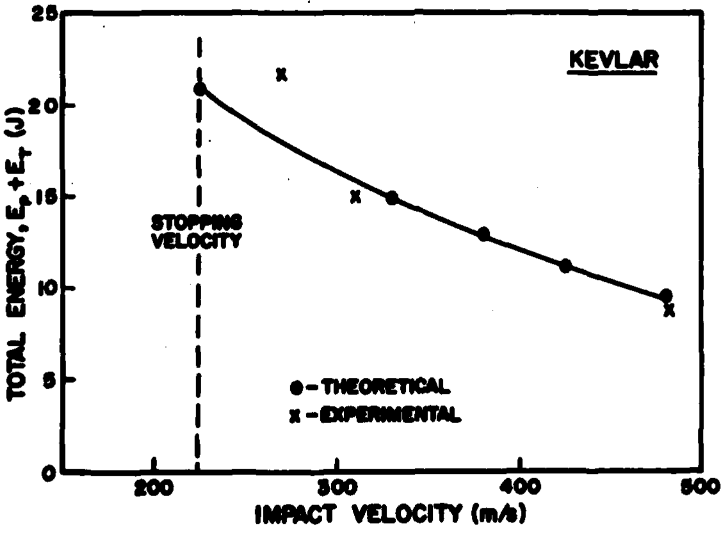


Figure 12. Comparison of Experimental and Theoretical Energy Absorption for Kevlar.

indicates that the tensile and transverse response modes predicted by equations 1 and 2 were active and the Kevlar achieved its maximum potential.

In the case of nylon (Figure 13), the theoretical energies were actually greater than those computed for Kevlar due to its superior yarn energy absorption potential. The experimental results however showed the actual performance of nylon to be in the range of 25-50% of potential. This inefficiency is attributed to the disruption of the idealized response pattern due to the influence of heat.

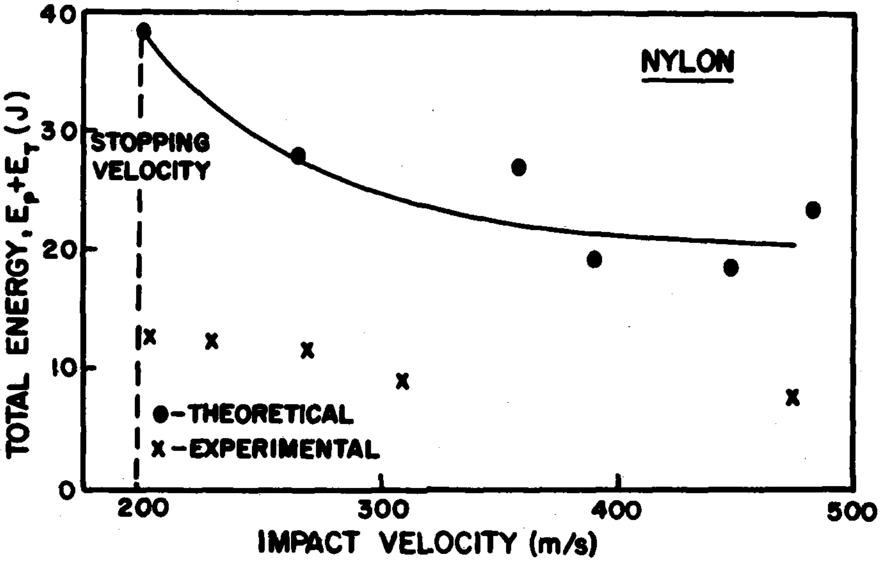


Figure 13. Comparison of Experimental and Theoretical Energy Absorption for Nylon

CONCLUSIONS

The energy from a ballistic projectile incident on a fabric target is dissipated through two distinct modes, namely energy absorbed within the fabric plane and energy exerted in displacing the fabric transversely.

Maximum ballistic resistance is achieved when the projectile energy dissipated within the fabric plane is absorbed through tensile resistance of the yarns. When this response mechanism is active, the time for projectile penetration is maximized, resulting in increased material involvement and increased energy absorption through both modes.

Performance models based on idealized tensile response and practical fabric parameters can be used to predict maximum expected energy absorption. Deviation from predicted performance is indicative of non-tensile failure responses such as melting and shear which severely reduce ballistic efficiency. The scanning electron microscope is a valuable device for verification of such failure patterns.

Strain wave velocity is the most important parameter influencing ballistic performance. Rapid propagation of stress away from the point of impact incorporates more material into the resistive process. It also distributes the energy such that the energy/mass ratio is kept below levels at which shear or heat failures may otherwise occur. Strain wave velocity is significantly affected by fabric construction. Potential for improved ballistic performance exists through loose fabric weave design and reduced yarn crimp.

REFERENCES

(1) Hearle, U. of Manchester Inst. of Sci & Tech., 1st Periodic Report, Contract No. DAJA37-69-C-0532.

(2) Roylance, D., and Wang, S., US Army R&D Labs, TR-80/021 (1979).

(3) Figucia, F., US Army Natick R&D Labs, TR-80/008, (1979).

(4) Figucia, F., Army Science Conference Proceedings, Vol. II, p. 29, (1980).

(5) Koza, W., US Army Natick R&D Labs, IPL-249, publication pending, (1982).

(6) Figucia, F., Weiner, L., and Laible, R., Polymer Eng & Science, Vol. II, No. 4, (1971).

(7) Unpublished data, US Army Natick R&D Labs, Notebooks No. 5789,(1970) and 6098, (1972).

(8) Dogliotti, E.C., and Yelland, W.E.C., High Speed Testing, Vol. IV, Wiley & Sons, (1963).

(9) von Karmen, T. and Duwez, P., J. Appl. Phys., Vol. 21, p.987, (1950).

(10) Petterson, E.R., Stewart, G.M., Odell, F.A., and Maheux, R. C., Textile Res. J., Vol. 30, p. 411, (1960).

(11) Fenstermarker, C.A. and Smith, J.E., Appl. Poly. Symp., Vol. I, p. 125, (1965).

A NOVEL BEAM BUNCHING CONCEPT FOR MILLIMETER WAVE TUBES

*PAUL FISCHER, MR.
CALVIN BATES, MR.
JOSEPH HARTLEY, MR.
US ARMY ELECTRONICS TECHNOLOGY AND DEVICES LABORATORY
FORT MONMOUTH, N.J. 07703

INTRODUCTION

It can be demonstrated(1) that under suitable boundary conditions, a post-loaded reentrant RF cavity configuration such as that shown in Figure 1 can produce an oscillating electron sheet in its gap region when the cavity is excited by RF energy.

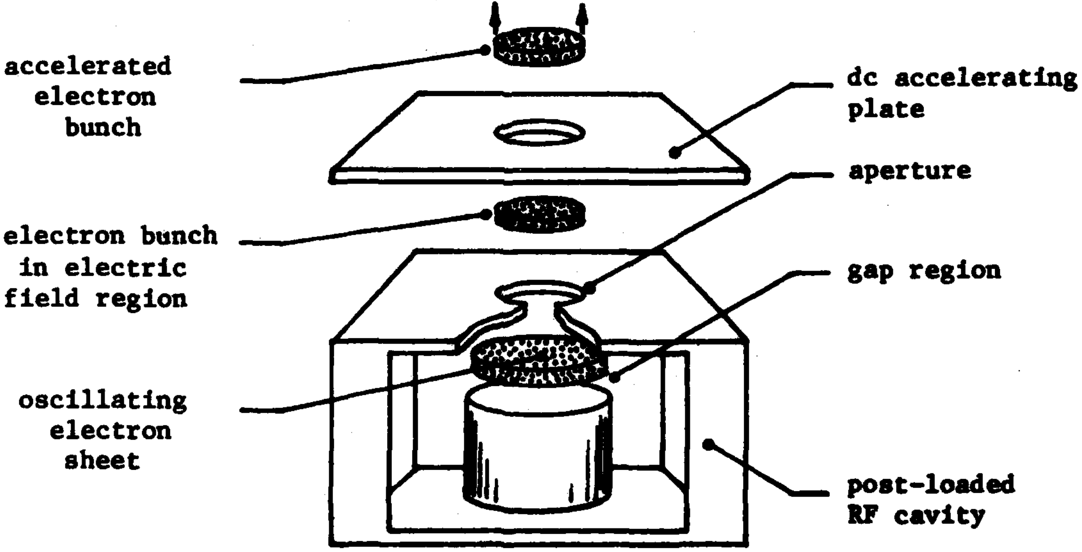

accelerated electron bunch
dc accelerating plate
aperture
electron bunch in electric field region
gap region
oscillating electron sheet
post-loaded RF cavity

Figure 1. RF Electron Gun with dc Acceleration Region

The electron sheet is highly bunched and oscillates back and forth in the gap region at the fundamental frequency of the RF excitation. An aperture in the upper surface of the gap permits electron bunches from the oscillating sheet to escape from the cavity. One electron bunch is produced every RF cycle and each bunch (when viewed from the plane of the aperture) represents a pulse of unidirectional RF current. Thus, RF energy which excites the cavity is converted into electron beam energy and the cavity configuracion is, in effect, a non-thermionic source of pre-bunched electrons rich in RF harmonics. If the electron bunches emerging from the gap aperture are projected through a constant electric field region as shown in Figure 1, they will gain kinetic energy, which by virtue of the sharp, discrete, nature of the bunches can be converted into RF energy when the bunches are coupled to the fields of an RF output structure.

MATHEMATICAL MODEL AND ANALYSIS

The physical boundary conditions necessary to generate and control oscillating electron sheets in the cavity RF gap region can best be shown analytically with the aid of Figure 2 which is an idealized geometrical version of the RF gap and dc acceleration region of the gun.

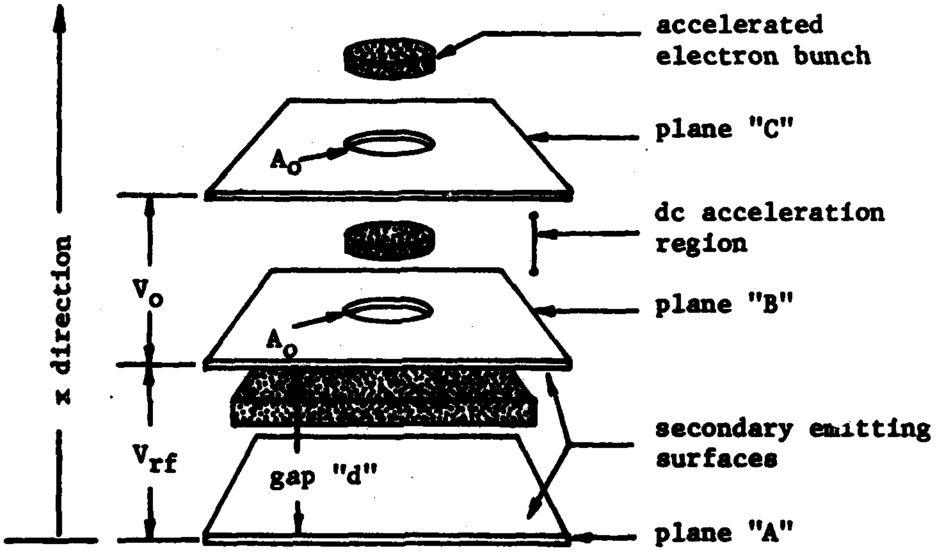


Figure 2. Analytical Model of RF Electron Gun

The RF gap region is modeled by two parallel planes, A, and B, with spacing, d, whose surfaces are capable of emitting secondary electrons when bombarded with primary electrons of sufficient energy. An RF voltage V_{rf}, is applied across planes A and B to generate and drive the electron sheet. The electron bunches are permitted to escape through an aperture of area, A_o, into a constant accelerating voltage region, V_o, bounded by planes B and C where the field imparts kinetic energy to the bunches. The accelerated electron bunches then pass through a second aperture of area, A_o, after which they can be utilized by an RF structure to convert their kinetic energy into RF output energy. An initial electron with velocity, v_o, starts out at plane A ($x=0$) and moves toward plane B whenever the RF field is in the proper direction. The acceleration of the electron at any instant of time, t, is given (in M.K.S. Units) by,

$$\frac{d^2x}{dt^2} = \frac{e}{m} \frac{V_{rf}}{d} \sin(\omega t + \phi_0) \qquad \{1\}$$

where ϕ_o is the phase of the RF field when $t = 0$, $x = 0$; ω is the angular frequency of the RF field, and e/m is the charge-to-mass ratio of the electron. The velocity, v, of the electron at any position in the RF gap is obtained by integrating Equation {1} once with respect to time and is given by,

$$v = \frac{dx}{dt} = v_o + \frac{e}{m} \frac{V_{rf}}{d\omega} \{\cos \phi_0 - \cos(\omega t + \phi_0)\} \qquad \{2\}$$

Integrating once again with respect to time, one obtains the expression for the instantaneous electron position as follows,

$$x = (v_o + \frac{e}{m} \frac{V_{rf}}{d\omega} \cos \phi_0)t + \frac{e}{m} \frac{V_{rf}}{d\omega^2} \{\sin \phi_0 - \sin(\omega t + \phi_o)\} \qquad \{3\}$$

Equations {1} to {3} contain implicit information needed to determine the conditions under which a planar RF gap region can generate and support an electron sheet which oscillates between the opposing faces of the gap at the frequency of the drive source as shown in Figures 1 and 2. In conjunction with the equations, one must use an empirically derived secondary emission curve (Figure 3) for the particular secondary emitters being employed on the gap surfaces. The figure illustrates a typical secondary emission curve in which primary electrons have sufficient energy to liberate secondaries whenever their impact energy in electron volts (ev) specified numerically by V_{imp}, is between V_I and V_{II}. V_I and V_{II} are the unity crossover points of the secondary emission ratio curve (ratio of secondary to normally incident primary electrons). In order for the sheet to exist, electrons within the sheet must traverse the gap distance d, (Figure 2) within, or very near one-half cycle of the drive oscillation frequency. When the sheet arrives at plane B, electrons within the sheet must have energies between V_I and V_{II} when they strike, so that they can produce secondary electrons to continue the process when the RF field reverses to draw the electrons back to plane A in the next half-cycle. Expressed mathematically, the transit time and impact boundary conditions are respectively,

$$t = \pi/\omega \qquad \{4\}$$

$$V_I < V_{imp} < V_{II} \qquad \{5\}$$

$$\text{for} \quad x = d \qquad \{6\}$$

$$\text{and} \quad v_o = 0 \qquad \{7\}$$

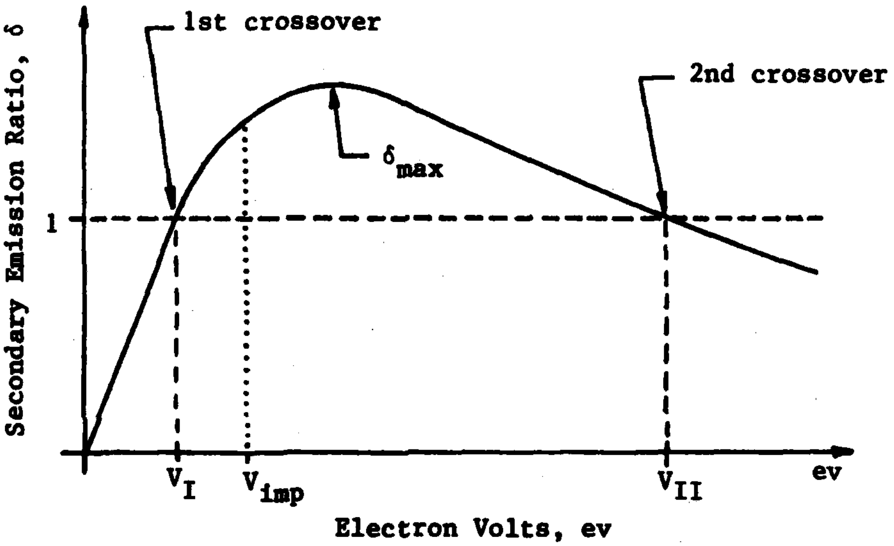


Figure 3. Generalized Secondary Emission Curve

Using Equation {4} through Equation {7} in conjunction with Equations {2} and {3} , yields, the gap dimension d, consistent with the required transit time condition shown in Equation {4}.

$$d = \sqrt{\frac{e}{m} \frac{V_{rf}}{\omega^2} \pi} \qquad \{8\}$$

The impact voltage (V_{imp}) corresponding to that gap is given by,

$$V_{imp} = \frac{2}{\pi} V_{rf} \qquad \{9\}$$

V_{imp} can be chosen independently from the experimental secondary emission curve (Figure 3); in practice it is preferable to pick V_{imp} at the peak, (δ_{max}), of the curve so that the maximum number of secondary electrons are obtained at impact each half-cycle. Once V_{imp} is specified, V_{rf} is determined by Equation {9}. The required V_{rf} is a function of the cavity (Figure 1) gap shunt impedance R_{sh}, and input drive power P_d, where,

$$P_d = \frac{V_{rf}^2}{R_{sh}} \qquad \{10\}$$

V_{rf}, therefore, can be obtained by suitable cavity design for R_{sh} or adjustment of the drive power, P_d or both. Since the shunt impedance and drive power can be chosen independently of frequency (2), V_{rf} is independent of frequency. Furthermore, the physical phenomenon of electron bunching is not dependent on frequency, but is determined by the phase of the electron leaving one surface, and the phase and energy with which it strikes the opposite surface. The electron bunching process occurs as a result of the phase and energy selection of electrons existing in the gap region. Thus, a scaled experimental model (for example 5 Gigahertz (GHz)) can be designed and built to verify the foregoing analysis for millimeter wave lengths. Once the shunt impedance and drive power are specified to produce the required RF voltage (V_{rf}), the gap spacing (d) is determined by Equation {8} for any desired drive frequency, ω.

The cyclic electron build-up and bunching process described in the foregoing analysis is a special case of the secondary electron resonance phenomenon (also called multipactor) extensively described in the scientific literature (see, for example, Reference 3). Equations {8} and {9} which specify the gap dimension and the corresponding electron impact energy (V_{imp}) were derived for an initial electron velocity and departure time of zero. In actuality, electrons are emitted over a range of departure times (departure angle) and arrive at the opposite gap face over a range of arrival times (conduction angle). The departure and arrival times are normally clustered about $t = 0$ and $t = \pi/\omega$, respectively, for practical cases. In addition, the emitted electrons have initial velocities corresponding to energies of 0 to 30 ev (4) with the maximum number of electrons emitted at approximately 2-5 ev. Iterative calculations using Newton's method (5), for Equations {2} and {3}, show that electrons having initial energies of up to 5 ev, will have departure and arrival times which cluster about $t = 0$ and $t = \pi/\omega$ respectively. The increased initial velocities will cause corresponding increases in departure and conduction angles. Experimental devices to be discussed in this report employ a gap spacing of 0.020 inches (0.508mm), with surfaces of either silver or copper (see Reference 4 for experimental secondary emission curves). For this gap, calculations yield values of approximately 30 degrees for the departure and conduction angles.

For a gap which can support secondary electron resonance, the electron sheet builds up rapidly once RF power is applied (normally in less than 10 cycles) until the space charge in the sheet becomes bounded by the perveance of the gap (6). Perveance in this case defines the maximum peak RF current I_p (in amperes) obtainable through the aperture at plane B (Figure 2). The peak RF current is represented by rectangular pulses (solid lines) as shown in Figure 4.

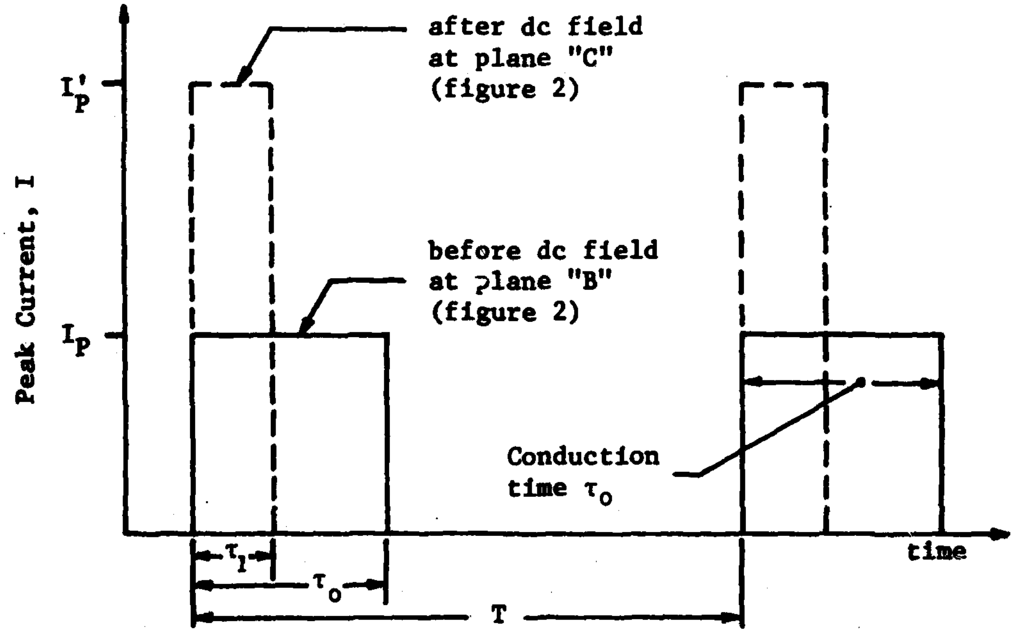

after dc field at plane "C" (figure 2)
I'_p
Peak Current, I
before dc field at plane "B" (figure 2)
I_p
Conduction time τ_o
time
τ_1
τ_o
T

Figure 4. Peak Currents Before and After Acceleration of Electron Bunches

It should be noted here that the peak RF currents and powers represented in Figures 4 and 5 occur at each RF cycle, and should not be confused with the conventional terminology which defines peak current and peak power produced by power bursts (pulses) containing many RF cycles. The peak RF current is given by,

$$I_p = 2.334(10^{-6}) \frac{A_o}{d^2} V_{rf}^{3/2} \qquad \{11\}$$

The peak instantaneous power in the bunch can also be represented by a rectangular pulse as shown in Figure 5 (solid line), with the peak power P_B in the bunch at plane B being given by,

$$P_B = (V_{imp})(I_p) \tag{12}$$

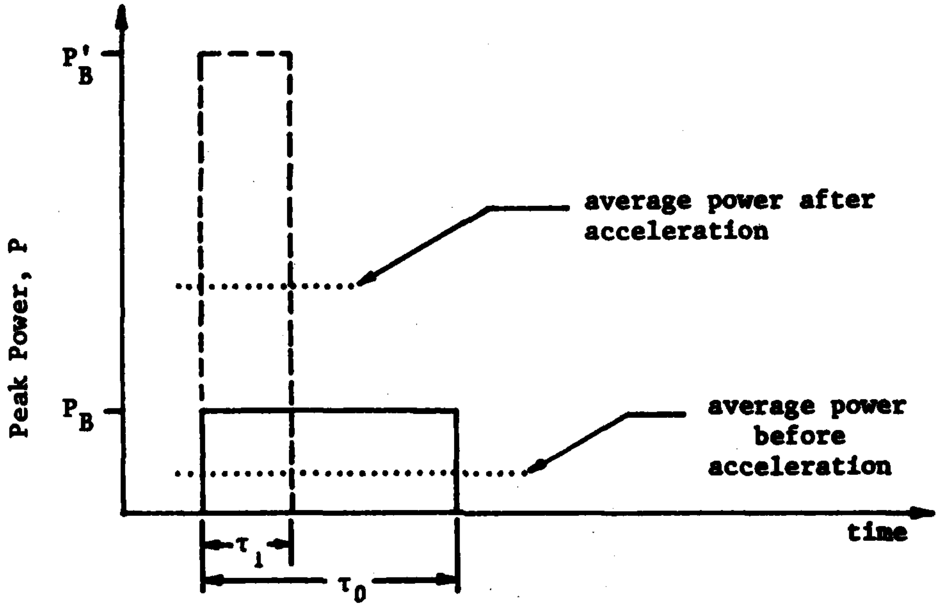

P'_B
P_B
Peak Power, P
average power after acceleration
average power before acceleration
time
τ_1
τ_0

Figure 5. Peak Powers Before and After Acceleration of Electron Bunches

After the bunch is accelerated through the dc acceleration region to gain energy, its peak power, P'_B, (dashed line) is given at plane C by,

$$P'_B = (G_p)(P_B) \tag{13}$$

where G_p is the peak power gain given by,

$$G_p = 1 + \frac{V_o}{V_{imp}}\left(1 + \sqrt{\frac{V_o}{V_{imp}}}\right) \tag{14}$$

and the corresponding peak current gain G_I, is,

$$G_I = 1 + \sqrt{\frac{V_o}{V_{imp}}} \tag{15}$$

The increased peak current, I'_p, and peak power, P'_B, at plane C are shown with dashed lines in Figures 4 and 5 respectively. For the experimental tube prototype to be discussed, $G_I = 3.2$ and $G_p = 12.2$ for $V_o = 1500$ volts and

V_{imp} = 300 volts (for copper). Thus, a substantial increase in the kinetic energy of the bunches can be obtained with a moderate accelerating dc voltage. The average power gain Gavp produced by the dc fields is,

$$G_{avp} = \frac{V_o}{V_{imp}} + \frac{1}{1 + \sqrt{\frac{V_o}{V_{imp}}}} \qquad \{16\}$$

and for the prototype tube Gavp = 5.3. Figure 5 also indicates that the average power is increased after acceleration (dotted lines).

The peak currents shown in Figure 4 can be represented by a Fourier series (7). For the purposes of this paper, we are interested only in the fundamental component of the RF current at plane C and its amplitude I_1, is given by,

$$I_1 = \frac{2}{\pi} I_P G_I \sin(\pi \frac{\tau_1}{T}) \qquad \{17\}$$

where τ_1, is the pulse width of the current at plane C, and T is the period of the fundamental RF drive. The ratio of the pulse width τ_1 at plane C, to the pulse width, τ_o, at plane B is,

$$\frac{\tau_1}{\tau_0} = \frac{1}{1 + \sqrt{\frac{V_o}{V_{imp}}}} \qquad \{18\}$$

and shows the effect of the dc accelerating field (V_o) on the pulse widths of instantaneous current and power. Neglecting RF circuit losses, one can show that the maximum RF power gain G_{rf} (with respect to the fundamental component) achieveable with the electron bunches is,

$$G_{rf} = 1 + \frac{V_o}{V_{imp}} \qquad \{19\}$$

and the maximum efficiency, η, is,

$$\eta = \frac{1}{1 + 2\frac{V_{imp}}{V_o}} \qquad \{20\}$$

Thus, the maximum limitations of RF power gain and efficiency (neglecting the RF losses in the RF input and output structures) with the gun/bunching configuration as shown in Figures 1 and 2, are functions of the impact voltage (V_{imp}) and the dc accelerating voltage (V_o). From Equations {19} and {20}, it follows that in order to optimize gain and efficiency, one must minimize V_{imp} and maximize V_o. V_{imp} in practice for presently useable secondary emitting materials (copper, silver, gold, platinum, magnesia-gold-cermets, etc.) is between 35 and 1000 ev. Power-supply/modulator, voltage breakdown, and electron de-bunching considerations affect the choice of V_o. In practice voltages as high as 25 kilovolts (kV) can be easily utilized to achieve very high gain and efficiency. However it can be shown that even voltages of approximately 5 kV will produce devices having useable gain and efficiency. For example, for copper, the peak δ in Figure 3 is approximately 1.25 at $V_{imp} = 500$ V. If one chooses $V_o = 5{,}000$ V, the maximum achievable power gain and efficiency are 10.4 dB and 83 percent respectively. Note that both maximum gain and efficiency are independent of frequency and improve with the dc accelerating voltage. In an actual device the power gain and efficiency are reduced somewhat by the RF losses in the input cavity and the RF output structure.

PRELIMINARY EXPERIMENTAL MODEL AND RESULTS

An experimental multipactor collector model shown in Figure 6 was designed and constructed at 5 GHz to demonstrate the RF to dc conversion features of the theoretical concept. Initial evaluations using a continuous wave (CW) input power level of 400 milliwatts (mW) produced current at the accelerator/collector electrode, and verified the 400 mW turn-on design value. The gap in this model was 0.020 inches (0.508 mm), and the secondary emitting surfaces were silver. The input power was increased to 1.5 watts to obtain sufficient RF drive to perform current collection experiments under stable conditions. The calculated current level for this design was approximately 8.5 mA for a 1.5 watt input, and the experimental value obtained was approximately 1 mA dc with a collector voltage of 100 V dc. The lower current obtained was most likely due to the RF

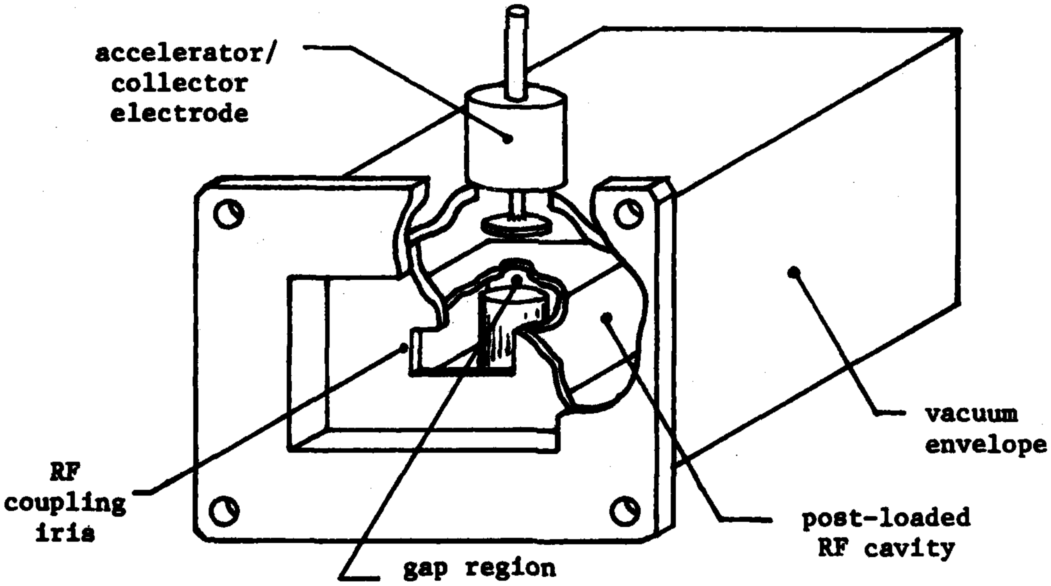


Figure 6. Experimental Multipactor Collector Model

losses in the cavity, and the large amount of outgassing noticed during operation which may have contaminated the secondary emitting surfaces. Pulsed drive tests were also performed, and Figure 7 shows the detected current pulse and detected RF input pulse. The duty was maintained at low level (<0.01) to minimize outgassing effects.

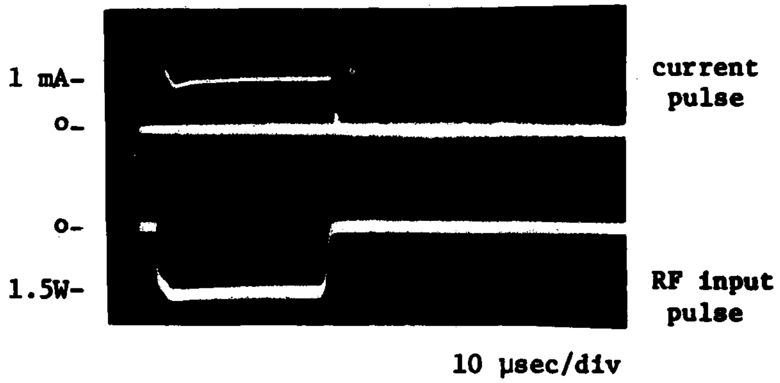


Figure 7. Detected Current Pulse and RF Input Pulse of Multipactor Collector Model

Again a peak current of approximately 1 mA was achieved with a peak power input of 1.5 watts. Note that the two pulses occur simultaneously showing that the turn-on and turn-off of the output current is controlled by the RF input pulse.

For further confirmation of the theory, the impact voltage (V_{imp}) of the electron bunches was measured by biasing the collector negative with respect to the aperture plane until current was cut-off. The cut-off occurred at approximately 300 volts which agrees within 10 percent of the calculated value of 289 volts.

RF PROTOTYPE DESIGN

A new tube prototype (8) was designed, based on the encouraging results obtained with the preliminary gun model. The new 5 GHz prototype is shown schematically in Figure 8. The tube consists of an RF input cavity which generates the electron sheet. An electron beam sieve, through which the electron bunches are driven, separates the input and output cavities. The RF output cavity ingeniously combines both the RF output interaction and the dc acceleration regions by virtue of an isolated accelerator electrode which is inserted into the hollow post of the output cavity. The input and output cavities are electrically discharged machined (EDM'ed) copper blocks as shown in Figure 9. Also shown are the accelerator electrode and the beam sieve plate.

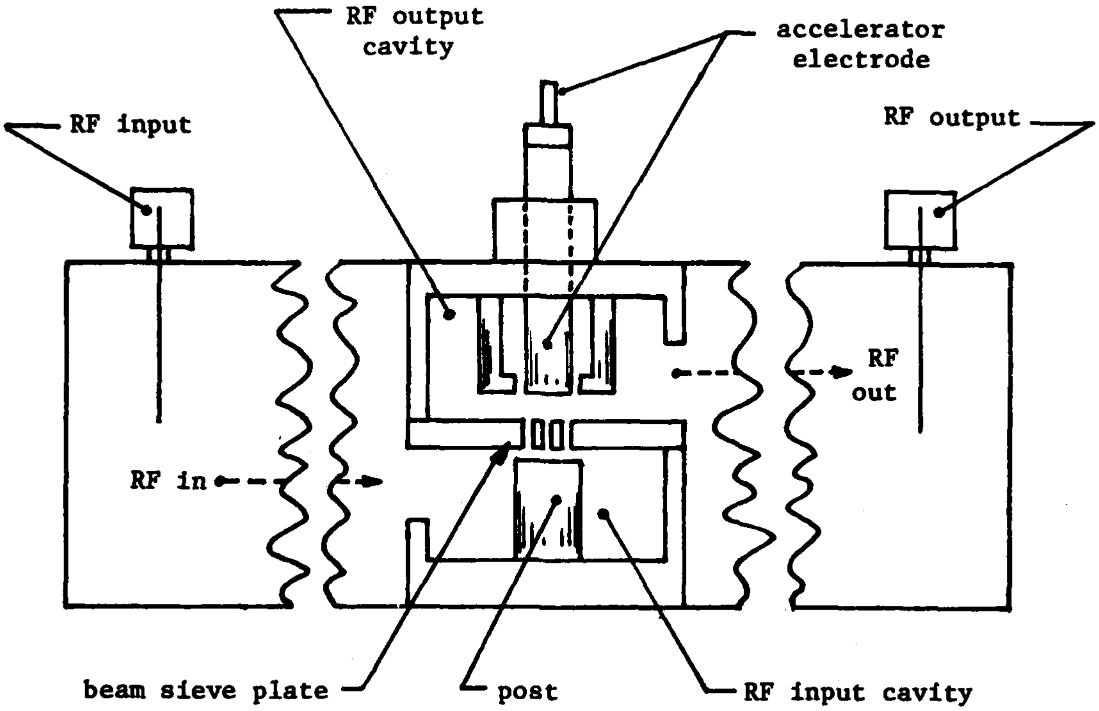

RF output cavity
accelerator electrode
RF input
RF output
RF out
RF in
beam sieve plate
post
RF input cavity

Figure 8. Internal Configuration of 5 GHz Prototype Tube

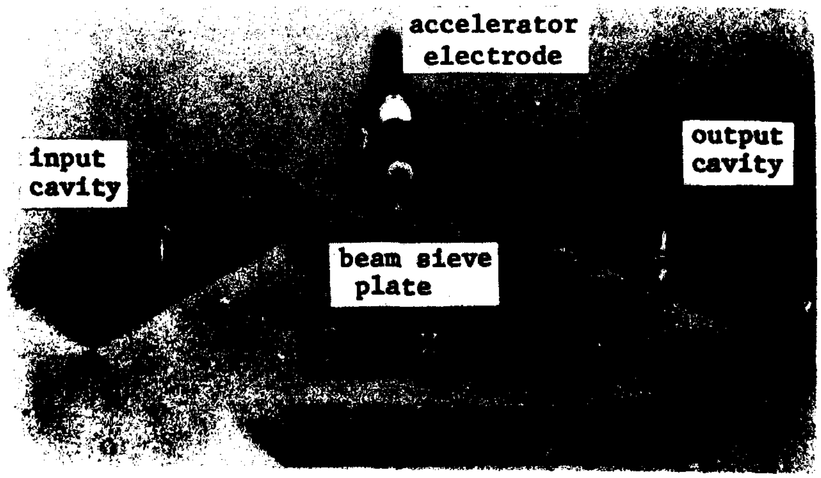

accelerator electrode
input cavity
output cavity
beam sieve plate

Figure 9. View of Internal Components of 5 GHz Prototype Tube

Figure 10 shows the input (lower), and output (upper) cavities stacked together with the accelerator electrode inserted in the output cavity. The RF input coupling iris and the post of the input cavity are visible. Cold tests of the RF subassembly were performed using standard network analyzer techniques to assure frequency coincidence of the input and output cavities with the beam-sieve plate and accelerator electrode in place. A single-hole beam-sieve plate which was described in the theory, provides the proper aperture for the electron bunches, however, the cold tests showed that the RF isolation between cavities was only 10 dB. In order to improve the RF isolation a multiple-aperture beam-sieve plate, using seven smaller holes, was fabricated and cold tests showed an improvement to 25 dB RF isolation. An RF isolation of at least 20 dB was considered important to separate the effects of the cavities during future hot tests, and to assure that the output power to be observed would be a function of the RF gain mechanism, and not a function of RF leakage between the cavities. Moreover, the multiple-aperture beam-sieve will be more effective than the single-aperture sieve in coupling the RF fields in the output cavity to the electron beam, since the RF fields generated between the post and metallic surfaces of the sieve will be in very close proximity to the electron beam. The multiple-aperture beam-sieve will also enhance the operation of the input cavity, by providing a more uniform RF field distribution in the gap region, than that attainable using a larger single-aperture sieve.

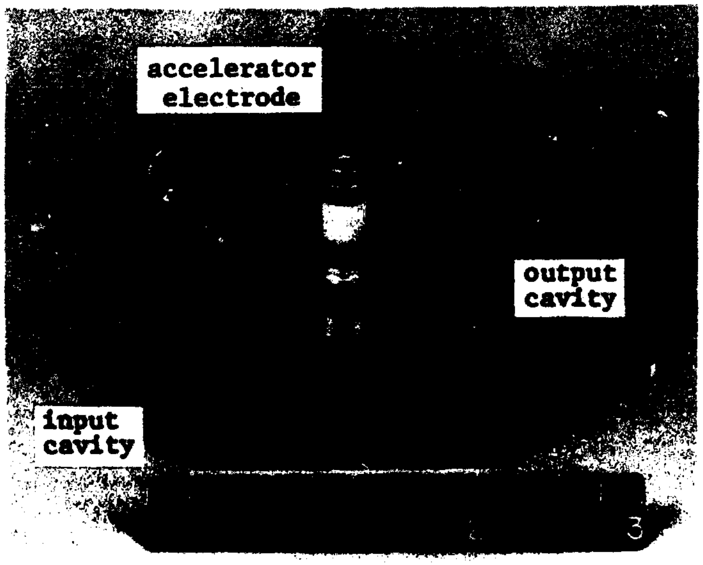


Figure 10. RF Subassembly for 5 GHz Prototype Tube

The RF subassembly is inserted into the center of the waveguide body, where it is held in place by the accelerator electrode. The complete prototype tube (Figure 11) uses standard RF vacuum connectors which are brazed to the waveguide body at the same time the accelerator electrode, vacuum fittings, and end plates are brazed. The tube will be capable of a bakeout temperature of 500°C.

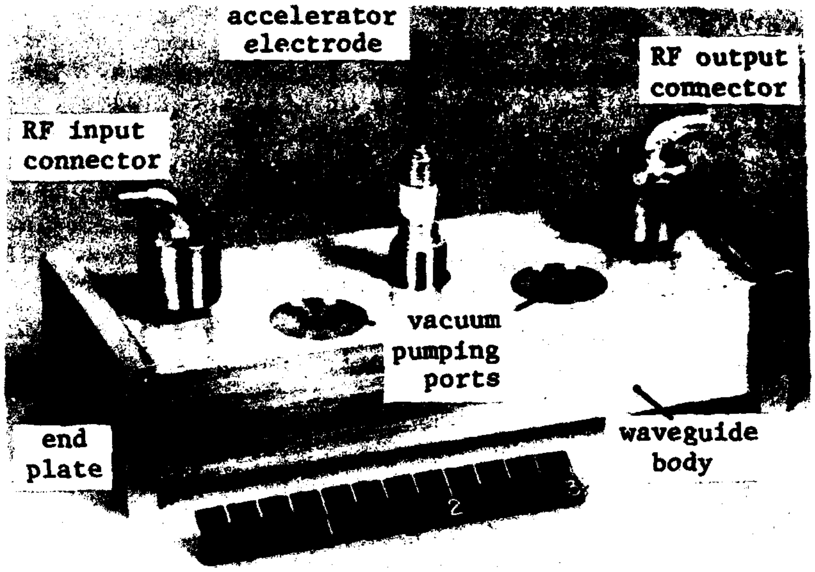

accelerator electrode
RF output connector
RF input connector
vacuum pumping ports
end plate
waveguide body
2
3

Figure 11. Complete Assembly of 5 GHz Prototype Tube

The design parameters for this tube which utilizes copper gap surfaces are listed below.

Frequency	5 GHz
Multipactor gap, (d)	0.020" (0.5mm)
Electron impact, voltage (V_{imp})	300 volts
Gap RF voltage	471 volts
Accelerator voltage (V_o)	1.5 kV
RF gain	8.2 dB
Efficiency	70%
Input power	1.5 watts
Output power	10 watts
Volume	15 cubic inches (250 cc)
Weight	1.5 pounds (0.7 kg)

This feasibility prototype has been designed and fabricated to demonstrate the formation of the electron bunches, and the gain mechanism generated by interaction with the RF fields. Future multi-stage models designed around the concept will demonstrate enhanced capabilities for specific Army system application.

CONCLUSIONS

A unique electron beam bunching concept has been presented, analyzed, and applied to the development of a novel tube approach. The analysis provides a design base for a new class of millimeter wave tubes. These tubes will be compact, lightweight, reliable, and low cost due to the simplicity of design and construction in comparison to conventional thermionic tubes with magnets. Furthermore, the analysis shows that useful RF gain and efficiency can be achieved at moderate accelerator voltages, which makes possible the use of simplified, low cost, lightweight, power supplies and modulators.

Theoretical considerations have been confirmed by evaluation of an experimental multipactor collector model. Furthermore, the results have demonstrated that the electron bunching can be generated and controlled by low RF drive (less than 1 watt), and that the currents generated are useful for new tube approaches. Although experiments were conducted at 5 GHz, the results are applicable for devices operating at millimeter wave frequencies.

A 5 GHz prototype tube design has been formulated, and component parts for the tube have been machined and cold-tested. The tube is presently being assembled for subsequent processing and evaluation. It is expected that this simplified, two-cavity design will demonstrate useful gain and efficiency, and provide the technology base for developing new tube prototypes for Army millimeter wave system applications.

REFERENCES

(1) C. Bates et. al., "Multipactor Electron Gun for Millimeter Wave Tubes," IEDM Technical Digest, pp 339-343, Dec 1981

(2) T. Moreno, Microwave Transmission Design Data, McGraw-Hill Book Co., Inc., First Edition, pp 228-230, 1948

(3) A. J. Hatch, "Electron Bunching in the Multipacting Mechanism of High-Frequency Discharge," Journal of Applied Physics, Vol. 32, No. 6, pp 1086-1092, June 1961

(4) H. Bruining, Physics and Applications of Secondary Electron Emission, Pergamon Press, Ltd., First Edition, pp 27-31, 104-107, 1954

(5) D. McCracken, FORTRAN With Engineering Applications, John Wiley & Sons, Inc., pp 44-52, 1967

(6) W. J. Gallagher, "The Multipactor Effect," IEEE Transactions on Nuclear Science, Vol. NS-26, No. 3, pp 4280-4282, June 1979

(7) S. Goldman, Frequency Analysis Modulation and Noise, McGraw-Hill, pp 25-27, 1948

(8) C. Bates, "Multipactor Beam Amplifier," Invention Disclosure, Docket No. 2447, filed May 1981

NEW VIEWPOINTS IN MASS FILTER DESIGN

*MELVIN H. FRIEDMAN, PhD
US ARMY MOBILITY EQUIPMENT R&D COMMAND, FORT BELVOIR, VA 22060

JOSEPH E. CAMPANA, PhD
US NAVAL RESEARCH LABORATORY, WASHINGTON, DC 20375

ALFRED L. YERGEY, PhD
NATIONAL INSTITUTES OF HEALTH, BETHESDA, MD 20205

1. INTRODUCTION

The quadrupole mass filter (QMF) is ubiquitous in military and civilian mass spectrometry laboratories. This instrument is used by chemists to determine the mass distribution associated with a sample. As illustrated in figure 1, a QMF has four pole pieces and each of these is carefully shaped to follow the equation of a hyperbola i.e. with the coordinate system shown the equations of the pole pieces are given by $(x^2 - y^2) = \pm r_o^2$. In a typical QMF, r_o and ℓ are .28 cm and 15 cm respectively (1). Detailed analysis shows (2) that mass filtering action is achieved by imposing a voltage difference $\phi = \phi_{DC} - \phi_{AC} \cos \omega t$ between adjacent electrodes. The driving frequency $\nu = (\omega/2\pi)$ is typically 2.5 MHz. The voltages ϕ_{DC} and ϕ_{AC} are refered to as the DC and AC components of the applied voltage respectively. The result of detailed analyses (1,2) is that the mass resolution of a QMF is determined by the ratio ϕ_{DC}/ϕ_{AC} while the absolute magnitude of the voltages ϕ_{DC} and ϕ_{AC} determine the mass of the ion which sucessfully traverses the length of the instrument. In the typical operation of the instrument (1) ϕ_{DC} and ϕ_{AC} are allowed to vary while their ratio is held constant and the number of ions which sucessfully traverse the filter are recorded.

The mass resolving power of a mass filter at mass m is defined by $m/\Delta m$ where Δm is a measure of the smallest difference in mass at mass m that can be detected by the instrument at a specified valley (e.g. 10% peak height) between peaks of equal heights. The transmission of a mass filter tuned to mass m is defined as the ratio of the number of ions with mass m injected into the mass filter to the number which arrive at the detector. It would be desirable for a QMF to have high resolving power and high transmission. Unfortunately, these are conflicting requirements and the improvement in one of these parameters results in a degradation of the other parameter.

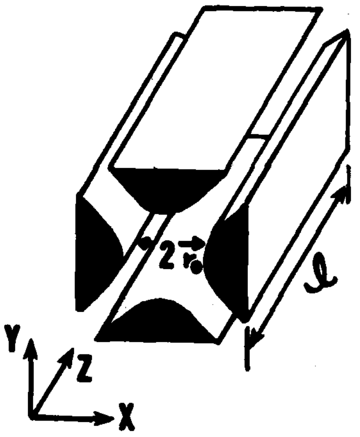


Figure 1. Electrode structure of the quadrupole and its relationship with the cartesian coordinate system. The coordinate system is oriented as shown but translated so that the origin is centered between the pole pieces as in figure 3. The ions, which move in a vacuum are introduced with a small velocity at $x=0$, $y=0$ in the z direction. By imposing a voltage difference with DC and AC components on adjacent electrodes the device is tuned so that only ions with the proper mass to charge ratio traverse the length of the filter where they are detected. Ions with a different mass to charge ratio collide with the electrode walls and are lost.

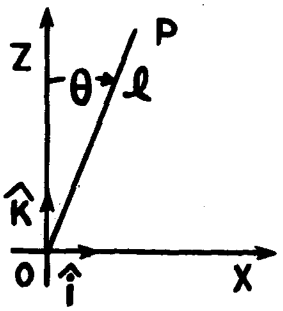


Figure 2. Inverted pendulum. The rigid rod OP of length ℓ is free to rotate in the x-z plane. The z-axis is oriented so it points away from the center of the earth and the support point O has an acceleration of $-b\omega^2 \cos \omega t$. For a high enough frequency ω the inverted pendulum has the remarkable property of being stable in a vertical position with its center-of mass above its support point.

The research reported here is motivated by the following question: Can the shape of the pole pieces or the time dependence of the applied potential difference be changed in such a way as to improve the resolution-transmission relationship? In part 2 of this paper, differential equations of ion motion are derived for geometries which generalize the popular quadrupole geometry. By integrating these equations of motion, the suitability of these geometries for mass spectrometry can be investigated. Should calculation indicate a particular geometry is suitable, the question of whether or not it yields improved performance can be investigated experimentally. Part 3 of this paper deals with the relationship between the QMF and the inverted pendulum. The inverted pendulum, which consists of a rigid rod free to rotate in a vertical plane and whose point-of-support is made to vibrate vertically has the remarkable property of being stable in a vertical position with its center of mass above its support point, providing the support point oscillates above a critical frequency. In part 3 of this paper it is shown that the differential equations for an ion traversing a QMF and the differential equations describing the motion of an inverted pendulum can be cast into the same form. In this way the inverted pendulum is shown to be a mechanical analog of the QMF. This aspect of the work is significant for three reasons: 1) The motion of an ion as it traverses a QMF cannot be physically seen. By observing the motion of an inverted pendulum one may better appreciate the subtleties involved in the way a quadrupole works and perhaps obtain a better sense of what is happening in this device. This may be of value to the person who tries to design a better mass filter. 2) The inverted pendulum has pedagogical value for explaining the operation of a QMF to students and has already been used for this purpose by one of the authors (A.L.Y). 3) By showing that the inverted pendulum and QMF are related subjects, it is possible that studies on one of them may shed light on the other.

The inverted pendulum has been described by several authors (3-10). Corben and Stehle (3) have modeled the inverted pendulum as a rigid rod while Landau and Lifshitz (4) have modeled it as a point mass attached to a rigid massless rod. Kapitza (5) has discussed the case of the pendulum when the amplitude of the sinusoidal oscillation of the suspension point is small in comparison to the length of the pendulum. He has also reported what appears to be the first operative pendulum with an oscillating support point (normal and inverted). Kalmus (6) has discussed the theory of the inverted pendulum with triangular excitation of the support point, and has reported an operative inverted compound pendulum. Mitchell (7) has applied a method of averaging to find the stability of an inverted pendulum for small amplitude high frequency oscillation, and Howe (8) has described a theory for the stability of an inverted pendulum driven by oscillations of a random nature. A square-wave model of the inverted pendulum has been recently described by Yorke (9). In a previous publication (10), Friedman et al. have observed that the inverted pendulum is a mechanical analog of the QMF.

Partial derivations of the differential equations for ion motion in a QMF are given by Campana (1) and also by Dawson (2). A derivation (starting from the Lorentz force law and the most general form of the Maxwell equations) for ion motion in multipole fields with a discussion of various approximations has been given by Friedman et al. (11).

2. ION MOTION IN MULTIPOLE FIELDS

a. Fundamental physics of ion motion. The force on an ion with charge e moving with velocity $\vec{v}$ through space in which there is an electric field $\vec{E}$ and a magnetic field $\vec{B}$ is given by the Lorentz force law

$$\vec{F} = e(\vec{E} + \vec{v} \times \vec{B}) \ . \qquad (1)$$

The differential equation which describes the motion of an ion through a field is obtained from Newton's law of motion

$$\vec{F} = m \frac{d^2\vec{r}}{dt^2} \quad . \tag{2}$$

It is necessary to compute the $\vec{E}$ and $\vec{B}$ fields caused by the time varying potential imposed on the pole pieces to find the motion of the ions from equations (1) and (2). These are computed from Maxwell's equations. If the charges on the ions and their associated current densities can be neglected for the purpose of computing the $\vec{E}$ and $\vec{B}$ fields acting on them, then in the vacuum region where the ions are moving Maxwell's equations are:

$$\vec{\nabla}\cdot\vec{E} = 0 \qquad \vec{\nabla}\times\vec{E} = -\frac{\partial\vec{B}}{\partial t}$$

$$\vec{\nabla}\cdot\vec{B} = 0 \qquad \vec{\nabla}\times\frac{\vec{B}}{\mu_\circ} = \frac{\partial}{\partial t}(\varepsilon_\circ\vec{E}) \quad . \tag{M1}$$

If the $\vec{E}$ and $\vec{B}$ fields are not changing with time then Maxwell's equations (M1) become:

$$\vec{\nabla}\cdot\vec{E} = 0 \qquad \vec{\nabla}\times\vec{E} = 0$$

$$\vec{\nabla}\cdot\vec{B} = 0 \qquad \vec{\nabla}\times\vec{B} = 0 \quad . \tag{M2}$$

The applicability of the two forms of Maxwell's equations (M1) and (M2) to ion motion in a mass filter merits discussion. Maxwell's equations (M1) are a good approximation in the vacuum region enclosed by the pole pieces of a mass filter. As pointed out in the introduction, the voltages on the electrodes change at radio-frequencies resulting in $\vec{B}$ and $\vec{E}$ fields which also change at the same frequencies. Therefore, it is not obvious that Maxwell's equations (M2) can be used for the purpose of computing ion motion. The use of equations (M2) for this purpose can be understood through the following discussion. Equations (M2) predict (12) that $B_\circ = E_\circ/c$ where $B_\circ$ and $E_\circ$ are the maximum amplitudes of the $\vec{B}$ and $\vec{E}$ fields and c is the speed of light. Thus the $\vec{B}$ field in the Lorentz force law (1) exerts a negligible force on the ion compared to the $\vec{E}$ field, unless the ion is moving at a speed comparable to the speed of light. Campana (1) has shown that the fastest ions in a quadrupole mass filter are moving slowly ($v/c < 10^{-4}$) compared to the speed of light, even under extreme operating conditions. Therefore, it is not necessary to consider the $\vec{B}$ fields any further for the purpose of computing the force on an ion. The Lorentz force law (1) reduces to $\vec{F}=e\vec{E}$, while equations (M2) become:

$$\vec{\nabla}\cdot\vec{E} = 0 \qquad \vec{\nabla}\times\vec{E} = 0 \quad . \tag{M3}$$

The second of these equations (M3) is valid only for an alternating potential which has a frequency ν low enough such that

$$\lambda = \frac{c}{\nu} >> \ell \quad . \tag{3}$$

That is, the wavelenth λ associated with the electromagnetic wave must be much greater than the length ℓ of the electrode structure. The necessity of equation (3) can be understood in the following way. If the wavelength does not satisfy equation (3) there is a possibility of standing waves (just as in the case of sound waves in an organ pipe) and these are not predicted by equations (M3). Typically, the length of the mass filter is less than one meter and the frequencies used (1) are well below 10 MHz. Therefore equation (3) is well satisfied and equations (M3) can be used to compute the ion trajectory.

An $\vec{E}$ field which satisfies equation (M3) can be found with the aid of two mathematical theorems. Stokes theorem, valid for any continuous, differentiable vector field $\vec{A}$, asserts that

$$\int\int_S (\vec{\nabla} \times \vec{A})\cdot d\vec{S} = \oint \vec{A}\cdot d\vec{r} \tag{4}$$

where $d\vec{r}$ is a vector on the circumference of the enclosed area S and whose direction is in the direction of integration. The vector $d\vec{S}$ has a magnitude equal to the area dS and its direction is normal to dS. The ambiguity in the two directions normal to dS is resolved by the right-hand rule, that is $d\vec{S}$ points in the direction of the thumb with the other fingers of the right hand pointing in the direction $d\vec{r}$. In equation (4), the integral on the right is over the entire circumference and the integral on the left is over the area S (not necessarily plane) enclosed by the circumference. Stokes theorem enables a two-dimensional integral over a surface which has the form on the left side of equation (4) to be evaluated by the simpler one-dimensional integral on the right-hand side of equation (4). Application of Stokes theorem to the second of equations (M3) implies

$$\oint \vec{E}\cdot d\vec{r} = 0 \tag{5}$$

over any closed path which defines a surface over which $\vec{\nabla} \times \vec{E} = 0$. The second mathematical theorem asserts that any vector field $\vec{E}$ which satisfies equation (5) around every closed path can be represented as the divergence of a scalar field $\phi(x,y,z)$, that is

$$\vec{E} = -\vec{\nabla}\phi(x,y,z) \quad . \tag{6}$$

Substituting $\vec{E}$ from equation (6) into the first of equations (M3) gives

$$\vec{\nabla}\cdot\vec{\nabla}\phi(x,y,z) = \nabla^2\phi(x,y,z) = 0 \ . \tag{7}$$

The prescription for finding the differential equations of ion motion for any configuration of pole pieces can now be given. Let $f_i(x,y,z)$

denote the equation for the surface S_i of conductor i. The differential equation of motion can be found by first obtaining a potential function $\phi(x,y,z)$ which satisfies Laplace's equation (7) and which also satisfies the boundary conditions

$$\phi(x,y,z)\Big|_{S_i} = \phi_i \quad . \tag{8}$$

Generally, it is difficult to find a function which satisfies equations (7) and (8). The function $\phi(x,y,z)$ can sometimes be found analytically using the method of images (12-14), or by complex variable theory (12-15), or by infinite series expansions (12-14,16) or it can be evaluated numerically using digital computers (17-19). Once the potential $\phi(x,y,z)$ is known, the differential equations of motion can be found using Newton's law

$$\vec{F} = m \frac{d^2\vec{r}}{dt^2} = -e\vec{\nabla}\phi(x,y,z) \quad . \tag{9}$$

As discussed earlier, this method is approximately valid for time-varying potentials

$$\phi_i = \phi_{DC} - \phi_{AC} \cos \omega t \tag{10}$$

providing the potential does not change too rapidly with time, i.e. subject to the limitations of equation (3). Physically, equation (10) corresponds to subjecting the electrodes to a voltage which has DC and AC components.

The remainder of this paper will consider only infinitely long conductors whose shape does not vary with z. Although real electrode structures must necessarily be of finite length, this approximation is made because it captures the essential features for the operation of electric RF devices while avoiding some of the mathematical complexities. With this approximation the potential for such a conductor configuration does not depend on z i.e. $\phi(x,y,z) = \phi(x,y)$ and hence the force in the z-direction is zero i.e. $m z = 0$. The equation of motion in the z-direction can be integrated directly

$$z = v_{z_0} t + z_0 \quad . \tag{11}$$

Here v_z and z_0 are the z-component of velocity and the z-coordinate respectively at t=0.

b. Method of obtaining potential function and equations of ion motion in multipole fields. The approach of the previous section was to find the potential function $\phi(x,y,z)$ for a specified set of conductors at potentials ϕ_i. An alternative approach is to find a function $U_n(x,y)$ which satisfies Laplace's equation (7) and from this function determine the equipotential surfaces of the pole pieces. If conductors with given applied potentials are fashioned which coincide with the equipotential surfaces, a potential function

$$\phi(x,y) = K_n U_n(x,y) \qquad (12)$$

can be found which satisfies Laplace's equation and the boundary conditions (8) by simply choosing the constant K_n properly. Physically, K_n is related to the dimension of the electrode structure and the applied electrode potentials; it is chosen such that $\phi(x,y)$ matches the known potentials at the equipotential surfaces.

Functions which satisfy Laplace's equation may be found by application of the theory of complex variables. If the complex variable $z = x + iy$ is raised to an integral power of n, then the result can be expressed as the sum of two functions, a real $U_n(x,y)$ and and imaginary $V_n(x,y)$ part:

$$(x + iy)^n = U_n(x,y) + iV_n(x,y) \quad . \qquad (13)$$

The integer n defines the order of the multipole field. From complex variable theory z^n is known to be analytic and so the function $U_n(x,y)$ satisfies (15) Laplace's equation (7).

The method described here for obtaining the potential function and the differential equation of motion can be summarised as follows: raise the complex variable z to an integer power n to find a potential function which satisfies Laplace's equation (7) and the boundary conditions (8); then ion motion is found from Newton's Law of motion (9).

c. *Application to some multipole geometries.* The method described in section 2b will now be applied to the hexapole geometry. The function z^n is evaluated for n=3 to obtain

$$U_3(x,y) = x^3 - 3xy^2 \quad .$$

Since $\nabla^2 U_3(x,y) = 0$, the function $U_3(x,y)$ is the basis for a possible potential function in charge free space. Its equipotential surfaces are illustrated in figure 4. Each of the pole pieces have the same shape and so the structure is unchanged under a 60 degree rotation. The pole piece lying on the positive x-axis is chosen to have positive applied potential $+\phi_\circ/2$ to comply with convention (2) and the adjacent electrodes are chosen to have potential with the same magnitude but opposite sign. The potential function $\phi(x,y)$ which satisfies Laplace's equation and the boundary conditions is

$$\phi(x,y) = \frac{\phi_\circ}{2r_\circ^3}(x^3 - 3xy^2) \qquad (16)$$

as is readily verified. With $\phi_\circ$ given by equation (10), Newton's law becomes:

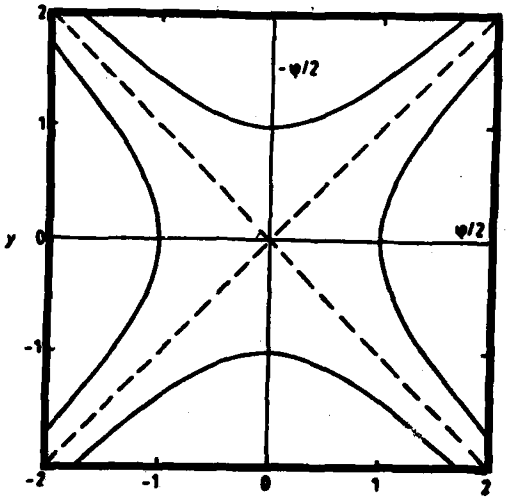

-φ/2
φ/2
y
2
1
0
-1
-2
-1
0
1
2

FIGURE 3. QUADRUPOLE GEOMETRY. EQUATION OF THE POLE PIECES:

$$\varphi(x,y)=(\varphi_0/2r_0^2)(x^2-y^2).$$

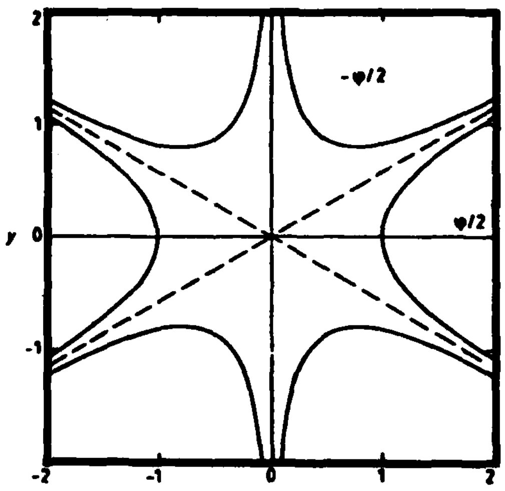

-φ/2
φ/2
y
2
1
0
-1
-2
-1
0
1
2

FIGURE 4. HEXAPOLE GEOMETRY. EQUATIONS OF THE POLE PIECES:

$$\varphi(x,y)=(\varphi_0/2r_0^3)(x^3-3xy^2).$$

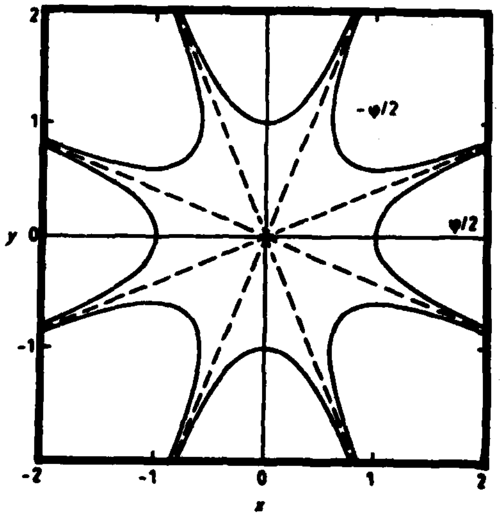

-φ/2
φ/2
y
x
2
1
0
-1
-2
-1
0
1
2

FIGURE 5. OCTAPOLE GEOMETRY. EQUATIONS OF THE POLE PIECES:

$$\varphi(x,y)=(\varphi_0/2r_0^4)(x^4-6x^2y^2+y^4).$$

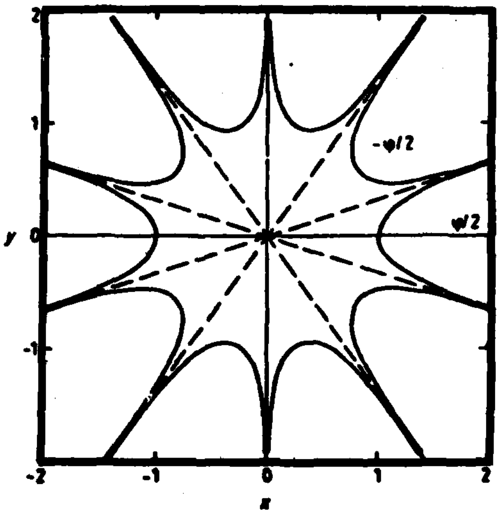

-φ/2
φ/2
y
x
2
1
0
-1
-2
-1
0
1
2

FIGURE 6. DECAPOLE GEOMETRY. EQUATIONS OF THE POLE PIECES:

$$\varphi(x,y)=(\varphi_0/2r_0^5)(x^5-10x^3y^2+5xy^4).$$

$$\begin{aligned} m\frac{d^2x}{dt^2} + \frac{3}{2}\frac{e}{r_\circ^3}(\phi_{DC} - \phi_{AC}\cos\omega t)(x^2 - y^2) &= 0 \\ m\frac{d^2y}{dt^2} - \frac{3e}{r_\circ^3}(\phi_{DC} - \phi_{AC}\cos\omega t)\,xy &= 0 \quad . \end{aligned} \tag*{(17)*}$$

The method just illustrated for n=3 determined the shape of the pole pieces (equation (16)) and the equations of ion motion (equation (17)) for the hexapole geometry. A similar procedure can be used for n any integer.

The geometry of the pole pieces when n=2 is the well-known quadrupole geometry illustrated in figures 1 and 3. The differential equation for ion motion in this case is:

$$m\frac{d^2x}{dt^2} + \frac{e}{r_\circ^2}(\phi_{DC} - \phi_{AC}\cos\omega t)\,x = 0 \tag{19a}$$

$$m\frac{d^2y}{dt^2} - \frac{e}{r_\circ^2}(\phi_{DC} - \phi_{AC}\cos\omega t)\,y = 0\ . \tag{19b}$$

The geometry of the pole pieces when n=4 is illustrated in figure 5. The differential equation for ion motion in this case is:

$$\begin{aligned} m\frac{d^2x}{dt^2} + \frac{2e}{r_\circ^4}(\phi_{DC} - \phi_{AC}\cos\omega t)(x^3 - 3xy^2) &= 0 \\ m\frac{d^2y}{dt^2} - \frac{2e}{r_\circ^4}(\phi_{DC} - \phi_{AC}\cos\omega t)(3x^2y - y^3) &= 0 \quad . \end{aligned} \tag*{(20)*}$$

The geometry of the pole pieces when n=5 is illustrated in figure 6. The differential equation for ion ion motion in this case is:

$$\begin{aligned} m\frac{d^2x}{dt^2} + \frac{5e}{2r_\circ^5}(\phi_{DC} - \phi_{AC}\cos\omega t)(x^4 - 6x^2y^2 + y^4) &= 0 \\ m\frac{d^2y}{dt^2} - \frac{10e}{r_\circ^5}(\phi_{DC} - \phi_{AC}\cos\omega t)(x^3y - xy^3) &= 0 \quad . \end{aligned} \tag*{(21)*}$$

The method illustrated in these examples can easily be used for finding the differential equations of ion motion for n>5.

3. THE INVERTED PENDULUM AND THE QUADRUPOLE MASS FILTER

a. Theory of the inverted pendulum. In this section the differential equation of a rigid pendulum with vibrating support point is derived. Figure 2 shows the pendulum which is a thin solid rod of length ℓ and mass m tilted at angle θ with respect to the positive z-axis and whose point of support is moving with speed $\dot{Z}$ in the positive z direction. The equation of motion for the pendulum is found by use of the Lagrangian function L where $L(q_k,\dot{q}_k) = T(q_k,\dot{q}_k) - V(q_k)$ and q_k, $\dot{q}_k$, T and V are the generalized coordinates, velocities, non-relativistic kinetic and potential energies respectively. The differential equation of motion is given by Lagrange's equation

$$\frac{d}{dt}\frac{\partial L(q_k,\dot{q}_k)}{\partial \dot{q}_k} - \frac{\partial L(q_k,\dot{q}_k)}{\partial q_k} = 0 \quad . \tag{22}$$

Consider the motion of the kth mass point m_k at distance ℓ_k from the support point. The velocity $\vec{v}_k$ of the mass point m_k is the vector sum of the velocity due to rotation and the velocity due to motion of the support point

$$\vec{v}_k = \frac{dx_k}{d\theta}\frac{d\theta}{dt}\,\hat{i} + \left(\frac{dz_k}{d\theta}\frac{d\theta}{dt} + \dot{Z}\right)\hat{k} \quad .$$

The geometrical relationships

$$x_k = \ell_k \sin\theta \qquad\qquad z_k = \ell_k \cos\theta$$

allow $\vec{v}_k$ to be expressed in terms of the generalized coordinate θ and generalized velocity $\dot{\theta}$

$$\vec{v}_k = \ell_k\,\dot{\theta}\cos\theta\,\hat{i} + (\dot{Z} - \ell_k\,\dot{\theta}\sin\theta)\,\hat{k} \quad . \tag{23}$$

Here $\hat{i}$ and $\hat{k}$ are unit vectors along the x and z axis respectively. Equation (23) and the definition of kinetic energy,

$$T = \frac{1}{2}\sum_k m_k v_k^2 = \frac{1}{2}\sum_k m_k\,\vec{v}_k\cdot\vec{v}_k \ ,$$

allow the kinetic energy to be expressed in terms of the generalized coordinates and velocities

$$T = \frac{1}{2}\left(\sum_k m_k \ell_k^2\right)\dot{\theta}^2 - \left(\sum_k m_k \ell_k\right)\dot{\theta}\,\dot{Z}\sin\theta + \frac{1}{2}\,m\,\dot{Z}^2 \quad . \tag{24}$$

As seen from figure 2, the potential energy V of the pendulum in terms of each mass point m_k is a decreasing function of θ given by

$$V = \left(\sum_k m_k \ell_k\right) g \cos\theta \tag{25}$$

where g is the acceleration of gravity. Equations (24) and (25) may be simplified by using the definitions for the moment of inertia I of a body rotating about the support point and the coordinate of the center of mass ℓ_c measured from the support point along the axis of the rod:

$$I = \sum_k m_k \ell_k^2 \qquad\qquad \ell_c = \frac{1}{m}\sum_k m_k \ell_k \quad .$$

If the mass is distributed symmetrically about its mid-point, then $\ell_c = \ell/2$ and Lagrange's equation can be written

$$L = \frac{1}{2} I \dot{\theta}^2 + \frac{1}{2} m \dot{Z}^2 - m \frac{\ell}{2} \dot{\theta} \dot{Z} \sin\theta - mg \frac{\ell}{2} \cos\theta \quad . \qquad (26)$$

With $q=\theta$, Lagrange's equation (22) applied to equation (26) yields the differential equation

$$I\ddot{\theta} - m \frac{\ell}{2} \ddot{Z} \sin\theta = mg \frac{\ell}{2} \sin\theta \quad . \qquad (27)$$

Two simplifications of equation (27) can be made. First, the moment of inertia for a uniform thin solid rod is easily computed (10)

$$I = m \frac{\ell^2}{3} \quad .$$

Second, the approximation $\sin\theta = \theta$ is made since it is anticipated that the oscillations of the pendulum will be stable near $\theta=0$ and so $\sin\theta$ will be small. Making these two substitutions, equation (27) becomes

$$\ddot{\theta} - \frac{3}{2\ell} (g + \ddot{Z}) \theta = 0 \quad . \qquad (28)$$

Equation (28), the principal result of section (3a), is seen to be reasonable if the cases $\ddot{Z}=0$, $\ddot{Z}=-g$ and $\ddot{Z}=+g$ are considered. For the case $\ddot{Z}=0$, equation (28) asserts that θ will grow exponentially with time constant $\tau = ((2\ell)/(3g))^{\frac{1}{2}}$. This is in agreement with the intuitive expectation that an inverted pendulum without oscillations is unstable. For the case $\ddot{Z}=-g$, the support point is moving down with a constant acceleration g. Equation (28) predicts that $\ddot{\theta}=0$ and this is in agreement with intuitive expectations since for a freely falling pendulum the effective forces acting to change the variable θ are expected to vanish. For the case $\ddot{Z}=+g$ the pendulum support point is accelerating upward with constant acceleration g. Equation (28) predicts that the motion for the variable θ is the same as that of a stationary pendulum sitting in a gravitational field of strength 2g and this too is in agreement with intuitive expectations.

b. The Mathieu equation. Subsequently it will be seen that the equation describing ion motion in a quadrupole and the equation describing the inverted pendulum can both be put in the cannonical form of the Mathieu equation. The purpose of this section is to describe this equation and its stability characteristics.

The canonical form of the Mathieu equation is

$$\frac{d^2u}{d\xi^2} + (a - 2q \cos 2\xi)u = 0 \quad . \qquad (29)$$

Here a and q are real constants, positive or negative, which are not functions of the dependent variable u or the independent variable ξ.

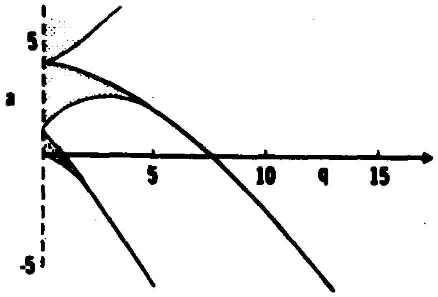


Figure 7. The a-q stability diagram. Equation (29) defines the parameters a and q which appear in the Mathieu equation. For a and q values which fall in the shaded area the value of u in equation 29 remains bounded as ξ increases independent of the initial conditions. Other values of a and q result in unbounded solutions.

Although the factor -2 multiplying q and the 2 in the argument of the cosine function might look strange, this is the standard for of this differential equation whose properties have been extensively studied.[22,23]

The solution of equation (29) will be a function of the form

$$u = u(\xi;\ a,q,u_0,\dot{u}_0) \quad . \tag{30}$$

Equation (30) expresses mathematically the observation that the function u depends on the independent variable ξ, on the parameters a and q, on the initial value of u denoted by u_0, and on the initial value of $du/d\xi$ denoted by $\dot{u}_0$. The function (30) evaluated for a particular value of a and q is a function of ξ, u_0 and $\dot{u}_0$. The result of detailed mathematical analysis (22) is that, depending on the values of the parameters a and q, the function u becomes unbounded or remains bounded as ξ increases and that these values of a and q are independent of the initial conditions u_0 and $\dot{u}_0$. Figure 7 shows a portion of a stability diagram in a-q space. The shaded area results in "stable" trajectories by which it is meant that u remains within finite bounds as ξ increases independent of initial conditions.

c. Analogy between inverted pendulum and quadrupole mass filter. In this section it is shown that both the inverted pendulum and the QMF can be expressed in the cannonical form of the Mathieu equation and the analogs will be identified.

In equation (28) $\ddot{Z}$ is the acceleration of the support point. If the support point is driven sinusoidally with frequency ω, then the origin of time can be chosen so that $Z = b \cos \omega t$. In that case, equation (28) becomes

$$\frac{d^2\theta}{dt^2} - \frac{3}{2\ell}\left(g - b\,\omega^2 \cos \omega t\right)\theta = 0 \quad . \tag{31}$$

Define four dimensionless parameters

$$\text{(a)}\quad u = \theta \qquad\qquad \text{(b)}\quad a = -\frac{6g}{\ell\omega^2}$$
$$\text{(c)}\quad \xi = \frac{\omega t}{2} \qquad\qquad \text{(d)}\quad q = -\frac{3b}{\ell} \tag{32}$$

and use these to replace the variables θ and t in equation (31) to get the cannonical form of the Mathieu equation (29).

Note that equations (19a) and (19b) are essentially the same aside from a change in sign and so without loss of generality only equation (19b) is considered in drawing the analogy with the inverted pendulum. Define four dimensionless parameters

$$\begin{array}{llll} \text{(a)} & u = y & \text{(b)} & a = -\dfrac{4e\phi_{DC}}{m\omega^2 r_0^2} \\ \text{(c)} & \xi = \dfrac{\omega t}{2} & \text{(d)} & q = -\dfrac{2e\phi_{AC}}{m\omega^2 r_0^2} \end{array} \qquad (33)$$

and use these to replace y and t in equation (19b) to get the cannonical form of the Mathieu equation (29).

From equations (32) and (33) the analogous quantities can be identified. Multiplying numerator and denominator of equation (32b) by m, the mass of the pendulum, and comparing equations (32b) and (33b), aside from some unimportant numerical factors, the quantity mg of the pendulum is analogous to $(e\phi_{DC})/r_0$ of the QMF; $m\ell$ of the pendulum is analogous to mr_0 of the QMF and ω of the pendulum is analogous to ω of the QMF. Multiplying numerator and denominator of equation (32d) by $m\omega^2$ and comparing equations (32d) and (33d), aside from some unimportant numerical factors, the quantity $bm\omega^2$ of the pendulum is analogous to $(e\phi_{AC})/r_0$ of the QMF; $m\ell$ of the pendulum is analogous to mr_0 of the QMF and ω of the pendulum is analogous to ω of the QMF. In these analogies, all the quantities associated with the pendulum have the same physical dimensions as the corresponding quantities associated with the QMF. Also, the RF source of the QMF and the oscillating source for the support point in the inverted pendulum perform analogous functions of producing a stability mechanisim for each system. The destabilizing terms of each system are g, the acceleration of gravity for the inverted pendulum, and ϕ_{DC}, the applied DC voltage of the QMF. The quantities which counter these destabilizing terms are $b\omega^2$, the acceleration amplitude for the support point of the inverted pendulum and ϕ_{AC}, the RF amplitude of the QMF.

4. CONCLUSIONS.

Explicit differential equations in closed form have been found for pole piece geometries which are generalizations of those used in the QMF. For arbitrarily shaped pole pieces, the solution of Laplace's equation cannot be written in closed form and, thus in general the ion differential equations of motion cannot be written in closed form. The geometries shown in figures 4, 5 and 6 are analytically useful in determining if other geometries might yield mass filters with improved resolution/transmission characteristics. Computer simulation of ion motion in these geometries is a subject for future research.

Although not reported in this paper, an inverted pendulum has been built (10) and it has been useful in emphasizing to students the non-intuitive nature of the mass filtering action that takes place in a QMF. In reference (10), it is shown conceptually that the inverted pendulum can be used to experimentally measure the boundaries between the stable and unstable regions in the a-q stability diagram. As a result, the inverted pendulum aids in understanding these diagrams which are crucial to understanding the operation of a QMF.

The principal new results of our research are: 1) new derivations of the equations which describe ion motion in a QMF, 2) the equations of pole pieces for hexapole, octupole and decapole geometries illustrated in figures 4, 5 and 6, 3) closed form differential equations of ion motion in these geometries (equations (17), (20) and (21)) and 4) the realization that the inverted pendulum is a mechanical analog of the QMF.

Today the quadrupole mass analyser is the most widely used mass analyser for low resolution applications. Typically, mass filters are used in laboratory environments where simplicity, compactness, economy, absence of magnetic fields, and/or the capability of fast scan rates (especially for chromatographic combinations) are important (1). Thus, an improvement in resolution/transmission qualities of electric RF mass filters would enable this widely used mass spectrometer to take a larger role in analytical laboratories and in basic research. Other design considerations besides resolution/transmission characteristics are important. If mass filters could be made which are smaller, more rugged, less susceptible to vibration with reduced power requirements, and more easily maintained, there might be greater application outside the laboratory. Possible applications include: surveying for natural resources, pin-pointing sources of industrial pollutants, and the detection of hazardous materials such as chemical agents and explosives. Eventually our work may be applied to the solution of these problems.

REFERENCES

1. J. E. Campana, Elementary Theory of the Quadrupole Mass Filter, Int. J. Mass Spectrom. Ion Phys. 33, 101 (1980).
2. P. H. Dawson (Ed), Quadrupole Mass Spectrometry and its Applications, Elsevier, Amsterdam (1976).
3. H. C. Corben and P. Stehle, Classical Mechanics, John Wiley and Sons, New York (1960).
4. L. D. Landau and E. M. Lifshitz, Mechanics, Pergamon Press, London (1960).
5. P. L. Kapitza, in D. ter Haar (Ed), Collected Papers of P. L. Kapitza, Pergamon Press, London (1965) pp 714, 726.
6. H. P. Kalmus, The Inverted Pendulum, Am. J. Phys. 38, 874 (1970).

7. R. Mitchell, Stability of the Inverted Pendulum Subjected to Almost Periodic and Stochastic Base Motion - An Application to the Method of Averaging, Int. J. Non-Linear Mech. 7, 101 (1974).
8. M. S. Howe, The Mean Square Stability of an Inverted Pendulum Subject to Random Parametric Excitation, J. Sound Vib. 32, 407 (1974).
9. E. C. Yorke, Square-Wave Model for a Pendulum With Oscillating Suspension, Am. J. Phys. 46, 285 (1978).
10. M. H. Friedman, J. E. Campana, L. Kelner, E. H. Seeliger and A. L. Yergey, The Inverted Pendulum: A Mechanical Analog of the Quadrupole Mass Filter, Am. J. Phys., in press.
11. M. H. Friedman, A. L. Yergey and J. E. Campana, Fundamentals of Ion Motion in Electric Radio Frequency Fields, J. Phys. E: Sci. Instr. 15, 53 (1982).
12. E. M. Pugh and E. W. Pugh, Principles of Electricity and Magnetisim, Addison-Wesley, Reading, MA (1960).
13. S. Ramo, J. R. Whinnery, T. van Duzer, Fields and Waves in Communication Electronics, John Wiley and Sons, New York (1965).
14. J. D. Jackson, Classical Electrodynamics, John Wiley, New York (1962).
15. R. V. Churchill, Complex Variables and Applications, McGraw-Hill, New York (1941).
16. R. V. Churchill, Fourier Series and Boundary Value Problems, McGraw-Hill, (1963).
17. K. J. Binns and P. J. Lawrenson, Analysis and Computation of Electric and Magnetic Field Problems, Macmillan, New York (1963).
18. M. G. Salvadori and M. L. Baron, Numerical Methods in Engineering, Prentice Hall, Englewood Cliffs, N.J. (1964).
19. V. Vemuri and W. J. Karplus, Digital Computer Treatment of Partial Differential Equations, Prentice Hall, Englewood Cliffs, N.J. (1981).
20. W. Paul and H. Steinwendel, A New Mass Spectrometer Without a Magnetic Field, Z. Naturforsch 8a, 448 (1953).
21. W. Paul. H. P. Reinhard, U. von Zahn, The Electric Mass Filter as Mass Spectrometer and Isotope Separator, Z. Phys. 152, 143 (1958).
22. N. W. McLachlan, Theory and Application of Mathieu Functions, Oxford University Press, London (1947).
23. M. Abramowitz and L. A. Stegun (Ed), Handbook of Mathematical Functions With Formulas, Graphs and Mathematical Tables, National Bureau of Standards, Applied Mathematics Series 55 (1964).

A NEW TYPE OF EM WAVE ABSORBING COATING

ARTHUR GAUSS, Jr., Ph.D.
BALLISTIC RESEARCH LABORATORY, ARRADCOM
ABERDEEN PROVING GROUND, MD 21005

Introduction

Research on radar anti-reflection coatings has been actively pursued both theoretically and experimentally since World War II. Interference, (1,2) dielectric (3), and magnetic-dielectric (4,5) absorptive coatings have been tried.

Recently a new type of radar absorbing chaff (RAC) has been suggested (6). Calculations done at BRL indicate that by compacting the RAC filaments a new type of radar absorbing coating (called RACO) may be possible. The filaments would be fixed in a solid binder having near unity dielectric constant. The primary mechanism for attenuation of the incident radiation by a RAC cloud or RACO layer is by resistive dissipation in the relatively long, thin chaff filaments. The tiny RAC-RACO filaments have length to diameter ratios on the order of one thousand with diameters of about 500Å. Filament separations in a RACO are about one-half to one-third of a filament length thus requiring the inclusion of dipole and perhaps higher order interactions. Each filament can be modeled simply as a series resistive-capacitive circuit, single filament theory will be discussed in Section II.

RACO's are divided into two classes, regular and random. Regular RACO's are regular arrays of parallel RAC filaments. Random RACO's are random collections of randomly oriented filaments. The resistive-capacitive nature of each filament enables us to express the dielectric constant of a RACO (regular or random) in terms of a simple relaxation mechanism which is mathematically equivalent to the Debye theory from solid state physics (7). The Debye formalism is readily adaptable to include dipole interactions. Regular RACO theory is analyzed in Section III, and reflectance data are calculated for a regular RACO over a metal substrate. In addition to the Debye theory the Onsager dielectric constant theory (8) is applied to random RACO's which are discussed in Section IV; again reflectance data are calculated.

In the analyses of both regular and random RACO's discussed above only a single relaxation time is considered. The relaxation time of a filament depends on its length and, in any actual RACO, there will be a distribution of filament lengths. In Section V the Cole and Cole theory (9) from solid state physics is shown to provide a convenient formalism for handling a distribution of filament relaxation times. Finally in Section VI the results are summarized and conclusions drawn.

II. Single Filament Theory

In this section we derive the dipole moment of a single filament. The dipole moment is a key ingredient in determining the dielectric constant of a RACO layer. Using the dielectric constant the reflectance of a RACO layer on top of a metal substrate can be determined.

A single filament can be considered as a series resistive-capacitance (R-C) circuit. An inductance term is not included since at the frequencies of interest here (10^{10} to 10^{11} Hz) it is not significant. The voltage (v) across the filament is written simply

$$v = R\frac{dq}{dt} + \frac{1}{C}q \qquad \text{(II.1)}$$

where q is the charge on the filament, R the resistance, and C the capacitance of the filament. Taking the external field parallel to the filament as $E_0 e^{i\omega t} . \ell$ we can write

$$v = E_0 \, e^{i\omega t} . \ell$$

where ℓ is the filament length. Since we wish to determine the dipole moment (p) we write

$$q = p/\ell \quad .$$

The resistance of the filament is given by

$$R = \frac{\ell}{\sigma A}$$

where σ is the filament conductivity and A is the cross sectional area. The capacitance is given by

$$C = \frac{A\varepsilon_0}{\ell L}$$

where ε_0 is the permittivity of free space and L is the depolarization factor, $L = \ln(\ell/d)/(\ell/d)^2$ where (ℓ/d) is the length to diameter ratio.

Substituting for v, q, R and C in (II.1) above we have an equation for the dipole moment:

$$E_0 e^{i\omega t} \cdot \ell = \frac{\ell}{\sigma A} \cdot \frac{\dot{p}}{\ell} + \frac{\ell L}{A\varepsilon_0} \cdot \frac{p}{\ell}$$

Taking $p = \bar{p} \; e^{i\omega t}$ and solving we obtain

$$\bar{p} = \frac{A\ell\varepsilon_0 \; \sigma}{(i\omega\varepsilon_0 + L\sigma)} \quad E_0 \qquad \text{(II.2)}$$

writing $A = \pi d^2/4$ we get

$$\bar{p} = \frac{\pi d^2\ell \; \varepsilon_0 \; \sigma}{4(i\omega\varepsilon_0 + \; \sigma L)} \quad E_0 \qquad \text{(II.3)}$$

which agrees with the expression obtained by Swinford (6) using a different procedure.

Now in a RACO layer where filament interactions become important the external field E_0 must be replaced by the local field E_L at a filament. Thus

$$\bar{p} = \frac{\pi \; d^2\ell \; \varepsilon_0 \; \sigma}{4(i\omega\varepsilon_0 + \sigma L)} \quad E_L \qquad \text{(II.4)}$$

III. Regular RACO's

A typical regular RACO is pictured in Fig. 1. A RACO is an artificial dielectric where the filaments play the role of atoms or molecules in a conventional dielectric. The general expression for the dielectric constant (ε_r) of any dielectric material (including artificial dielectrics) is

$$\varepsilon_r = 1 + \frac{P}{\varepsilon_0 \; E} \qquad \text{(III.1)}$$

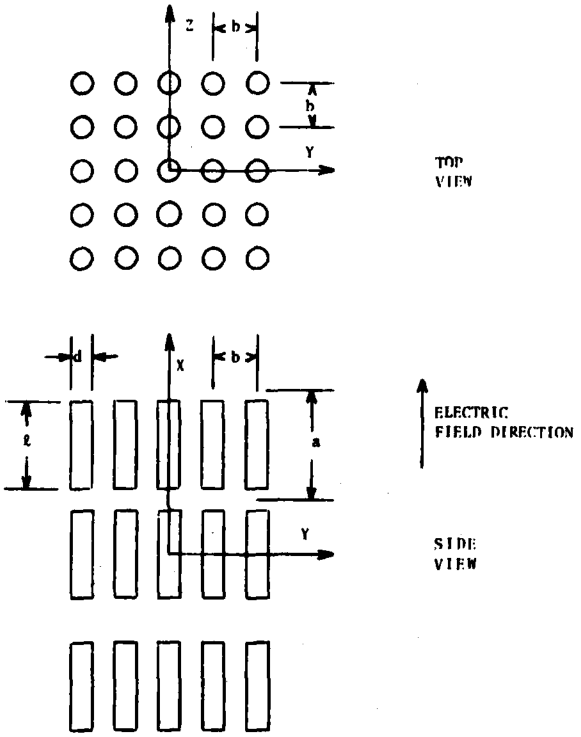


Figure 1. Regular RACO

where P is the polarization of the dielectric and E is the average field in the dielectric. For an artificial dielectric regular RACO

$$P = N\bar{p} \qquad \text{(III.2)}$$

where N is the number of filaments per unit volume and $\bar{p}$ (the dipole moment) is defined by eq. (II.4) repeated here

$$\bar{p} = \frac{\pi d^2 \ell \epsilon_0 \sigma}{4(i\omega\epsilon_0 + \sigma L)} E_L \qquad \text{(III.3)}$$

Expressing the polarization (P) in terms of the volume fraction f_v, $f_v = N(\pi/4)d^2\ell$ (for a regular array such as pictured in Fig. 1 we can also write $f_v = (\pi/4)\frac{1}{(\frac{a}{\ell})(\frac{b}{\ell})^2(\frac{\ell}{d})^2}$) , we can write

$$P = \frac{\varepsilon_0 \sigma f_v}{(i\omega\varepsilon_0 + \sigma L)} E_L \tag{III.4}$$

It can be shown that the average field E in the dielectric is related to the local field E_L by the following (10,11)

$$E = E_L - A \frac{P}{\varepsilon_0} \tag{III.5}$$

where A is the dipole interaction factor and depends on the array geometry of the RACO. Substituting (III.5) into (III.1) and using (III.4) we obtain

$$\varepsilon_r = 1 + \frac{\frac{x}{1-Ax}}{1 + i\omega\tau \left(\frac{1}{1-Ax}\right)} \tag{III.6}$$

where τ is the single filament relaxation time given by the product RC (see Section II) $= \varepsilon_0/\sigma L$, and $x = f_v/L$.

The above, (III.6), has the Debye form of the dielectric constant from solid state physics. The Debye form is (7)

$$\varepsilon_r = \varepsilon_{r\infty} + \frac{\varepsilon_{r0} - \varepsilon_{r\infty}}{1 + i\omega\tau_\varepsilon} \tag{III.7}$$

which matches (III.6) with the identifications

$$\varepsilon_{r\infty} = 1$$

$$\varepsilon_{r0} - \varepsilon_{r\infty} = \frac{x}{1-Ax}$$

and

$$\tau_\varepsilon = \tau\left(\frac{1}{1-Ax}\right)$$

For the purposes of reflectance calculations two versions of the dielectric constant expression will be used. One version is identical to

(III.6) and includes dipole interactions, the other version is obtained by setting A = 0 in (III.6) and corresponds to neglect of dipole interactions. Why after going to the trouble to include dipole interactions do we now neglect them? The answer lies in some experimental work (10) done in the 1950's on artificial dielectrics. The results of some of this work for a regular array of metal cylinders are shown in Fig. 2. These experiments were done on cylinders with ℓ/d ratios from 4 to 16, the frequency was a thousand hertz. Each cylinder was 0.062 in. in diameter. As can be seen from the figure for $0.7 < b/a < 1.6$ (b and a are defined in Fig. 1) the dipole approximate form of the dielectric constant is superior to the A = 0 form. However, for $b/a < 0.7$ (the situation encountered in most regular RACO's) the A = 0 case appears superior. This fortuitous agreement is thought to be caused by neglect of higher order interactions (quadrupole, octupole) that become important for very close filament separations.

The results of reflectance (reflection coefficient) calculations for a typical regular RACO are shown in Table I. The tabulated reflection

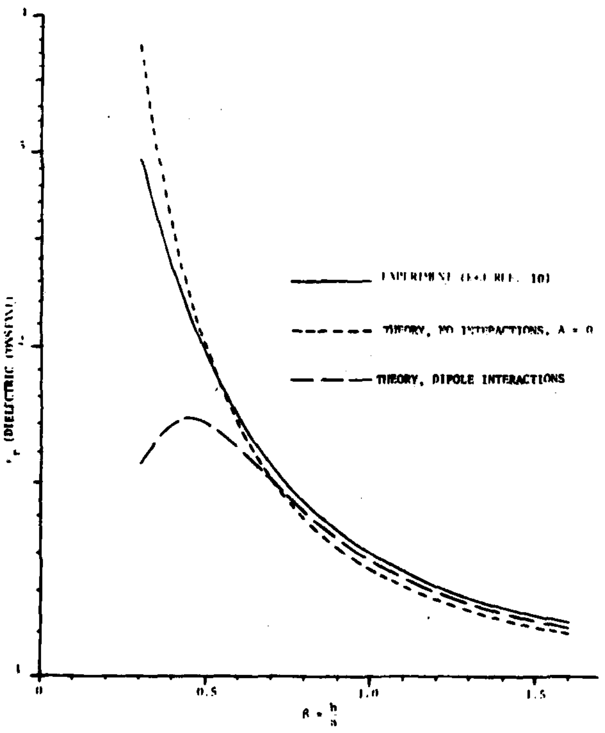


Figure 2. Comparison of theory and experiment for a regular array of parallel cylinders ($\frac{\ell}{d} = 4$, $\frac{\ell}{a} = 0.7$)

coefficient (R_f) is the fraction of incident intensity reflected for normally incident EM waves on a RACO over a metal substrate. The reflectance calculations were made using a computer program (12) developed at BRL which calculates reflectance at various frequencies of a layer or layers of dielectric material on top of a metal substrate. Additional comments on Table I are: (a) The reflectance rises rapidly for the A = 0 case below 10^{10} Hz and (b) the reflectance is very low over the whole range from 3 x 10^{10} Hz to 10^{11} Hz as shown by calculation of R_f at many intermediate points.

IV. Random RACO's

The theoretical development of the Debye form of the dielectric constant for random RACO's proceeds in a similar manner to the development for regular RACO's. The result for random RACO's is

---

TABLE I

Freq. (Hz)	R_f (Refl. Coeff.)	
	Debye A = - 1.696	Debye A = 0
10^{10}	0.412	0.0218
3×10^{10}	0.0248	0.0103
5×10^{10}	0.0059	0.0016
10^{11}	0.0051	0.00106

Filament Parameters

Conductivity	$\sigma \sim 10^5$ mho/m
Length/Diameter	$\ell/d \sim 1000$
Diameter	$d \sim 500$A

Array Parameters

Volume fraction $f_v = 5.65 \times 10^{-6}$ (computed using $b/\ell = 1/3$, $\ell/a = 4/5$ and ℓ/d above)

Thickness of array layer = 2 cm (on top of a perfect conductor)

$$\varepsilon_r = 1 + \frac{\frac{x'}{1-Ax'}}{1 + i\omega\tau\frac{1}{1-Ax'}} \qquad \text{(IV.1)}$$

where $x' = f_v/3L$ (the factor of 3 arises from the random orientation of the dipoles), $\tau = \varepsilon_0/\sigma L$ is the single filament relaxation time and A is the dipole interaction factor.

For the reflectance calculations we take two values of A in eq. (IV.1); namely $A = 0$ and $A = 1/3$. The first corresponds to the neglect of the dipole interactions between filaments whereas the second can be shown to represent a reasonable approximation for a random collection of aligned dipoles. It may be recalled that in most polar substances (and in random RACO's) the dipoles are not aligned and $A = 1/3$ is known to be a poor approximation (8). On the other hand some authors (13) believe $A = 0$ is the correct approximation for polar materials.

Another theoretical expression for the dielectric constant may be derived for subsequent use in reflectance calculations for random RACO's. This expression is similar to that used by Onsager (8) for random collections of polar molecules. In this theory a typical RACO filament represented by a point dipole is assumed to lie in a real spherical cavity in the artificial dielectric. The local electric field in the cavity is given by (14)

$$E_L = \frac{3\varepsilon_r}{2\varepsilon_r + 1} E \qquad \text{(IV.2)}$$

The dipoles are randomly oriented so the average dipole moment parallel to the field is $\overline{p'} = \overline{p}/3$. Hence using eq. (II.4) we obtain

$$\overline{p'} = \frac{A\ell\varepsilon_0 \sigma}{3(i\omega\varepsilon_0 + \sigma L)} E_L \qquad \text{(IV.3)}$$

Hence the polarization may be written as

$$P = \overline{p'}N \qquad \text{(IV.4)}$$

Substituting (IV.4), (IV.3) and (IV.2) into the basic equation

$$\varepsilon_r E = E + \frac{P}{\varepsilon_0} \quad , \quad \text{where } \varepsilon_r = \frac{\varepsilon}{\varepsilon_0} \quad ,$$

and defining

$$K^o = \frac{\sigma f_v}{3(i\omega\varepsilon_0 + \sigma L)} \quad ,$$

where $f_v = NA\ell = N\frac{\pi}{4}d^2\ell$ is the volume fraction, we obtain for the dielectric constant

$$\epsilon_r = \frac{3K^o + 1 \pm 3(K^{o2} + \frac{2}{3}K^o + 1)^{1/2}}{4} \qquad \text{(IV.5)}$$

The plus sign is to be taken in front of the radical since it is the only physical solution.

The results of reflectance (reflection coefficient) calculations for a typical random RACO are shown below in Table II. The tabulated reflection coefficient (R_f) is the fraction of incident intensity reflected for normally incident EM waves on a RACO over a metal substrate.

TABLE II

Freq (Hz)	R_f (Refl. Coeff.) Debye A = 1/3	 Debye A = 0	 Onsager
10^{10}	0.1203	0.2103	0.1840
3×10^{10}	0.0287	0.0299	0.0288
5×10^{10}	0.00803	0.00538	0.00784
10^{11}	0.000397	0.000425	0.000435

Filament Parameters (Same as for regular RACO - Table I)

Random RACO Parameters - Volume fraction $f_v = 3 \times 10^{-5}$. Thickness of RACO = 1 cm (on top of a perfect conductor).

The reflectance for the random RACO is very low over the whole range from 3×10^{10} Hz to 10^{11} Hz as shown by calculation of R_f at many intermediate points.

V. Cole & Cole Theory

In our analyses of regular and random RACO's discussed in detail in Sections III and IV only a single relaxation time was assumed. Given a constant filament conductivity this assumption implies filaments with a single length to diameter (ℓ/d) ratio. In any real RACO there will be a distribution of (ℓ/d)'s. To handle this situation the Cole & Cole theory (9) from solid state physics has been adapted to the RACO problem.

The Cole & Cole result (9) for the dielectric constant (ε_r) is written

$$\varepsilon_r = \varepsilon_{r\infty} + \frac{\varepsilon_{r0} - \varepsilon_{r\infty}}{1 + (i\omega\tau_0)^{1-\alpha}} , \tag{V.1}$$

where α is a parameter which determines the relaxation time distribution function and τ_0 is the central relaxation time of the distribution. For $\alpha = 0$ we obtain the Debye expression, eq. (III.7). The real part of the dielectric constant is

$$\varepsilon_r' = \varepsilon_{r\infty} + \frac{(\varepsilon_{r0} - \varepsilon_{r\infty})}{2} \left\{ 1 - \frac{\sinh\,(1-\alpha)s'}{\cosh\,(1-\alpha)\,s' + \sin\,(\frac{\pi}{2})\alpha} \right\} \tag{V.2}$$

where $s' = \ell n(\omega\tau_0)$. The imaginary part is given by

$$\varepsilon_r'' = \frac{\frac{(\varepsilon_{r0} - \varepsilon_{r\infty})}{2} \cos\,(\frac{\pi}{2}\alpha)}{\cosh\,(1-\alpha)s' + \sin\,(\frac{\pi}{2}\alpha)} \tag{V.3}$$

The Cole & Cole distribution function is given by

$$F(s) = \frac{1}{2\pi} \frac{\sin\,\alpha\pi}{\cosh\,(1-\alpha)s - \cos\,\alpha\pi} \tag{V.4}$$

where $s = \ell n(\tau/\tau_0)$. Fig. 3 shows the relaxation time distribution function (V.4) for various values of α. The $\alpha = 0$ distribution function is a spike corresponding to a single relaxation time, and the Debye theory. As α increases we generate progressively flatter distributions as shown by the figure.

Let us now consider a regular RACO for which $f_v = 2.5 \times 10^{-6}$. The filament parameters are: conductivity $\sigma = 10^5$ mho/m and length to diameter ratio $(\ell/d) = 1000$. Such a RACO has been demonstrated to be an effective absorber of cm-mm wavelength radiation. We will also assume the dipole interaction factor $A = 0$. For such a RACO (using the results of Section III) we have

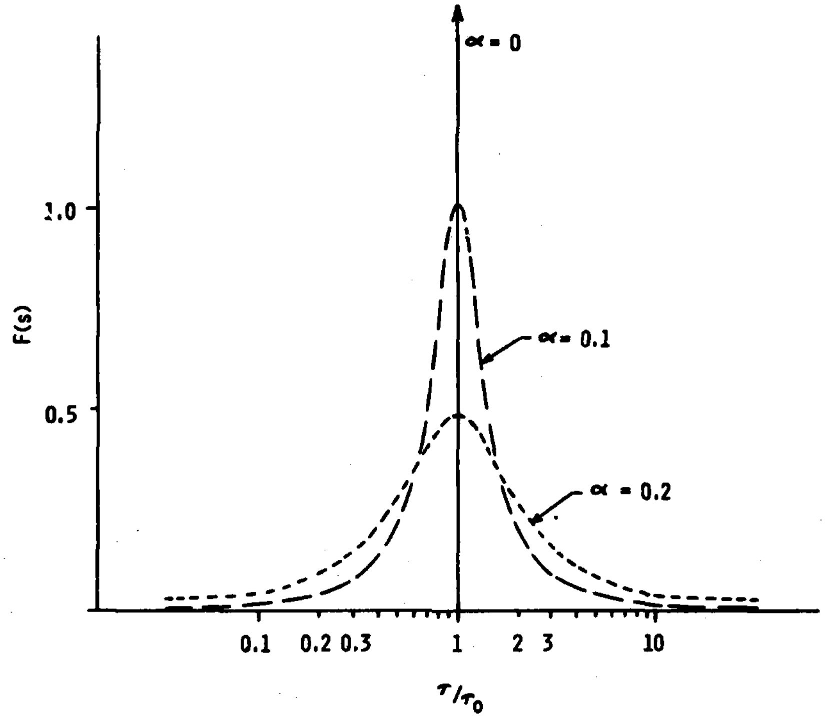


Figure 3. Relaxation time distribution function, Cole & Cole theory.

$$\varepsilon_{r0} - \varepsilon_{r\infty} = \frac{f_v}{L}$$

where the depolarization factor $L = \ln(\frac{\ell}{d})/(\frac{\ell}{d})^2$. Also

$$\tau_0 = \frac{\varepsilon_0}{\sigma L}$$

τ_0 and $(\varepsilon_{r0} - \varepsilon_{r\infty})$ can thus be easily determined. The real and imaginary parts of the dielectric constant can now be calculated. In Fig. 4 we plot the real part of the dielectric constant (eq. (V.2)) for the same α values used in the distribution function plots (Fig. 3). Note that even a fairly broad distribution corresponding to α = 0.2 does not change the real part of the dielectric constant significantly from the Debye α = 0 case. In Fig. 5 we plot the imaginary part of the dielectric constant. It too is not greatly changed. Experience with reflectance calculations for many different RACO's tells us that such modest changes in the dielectric

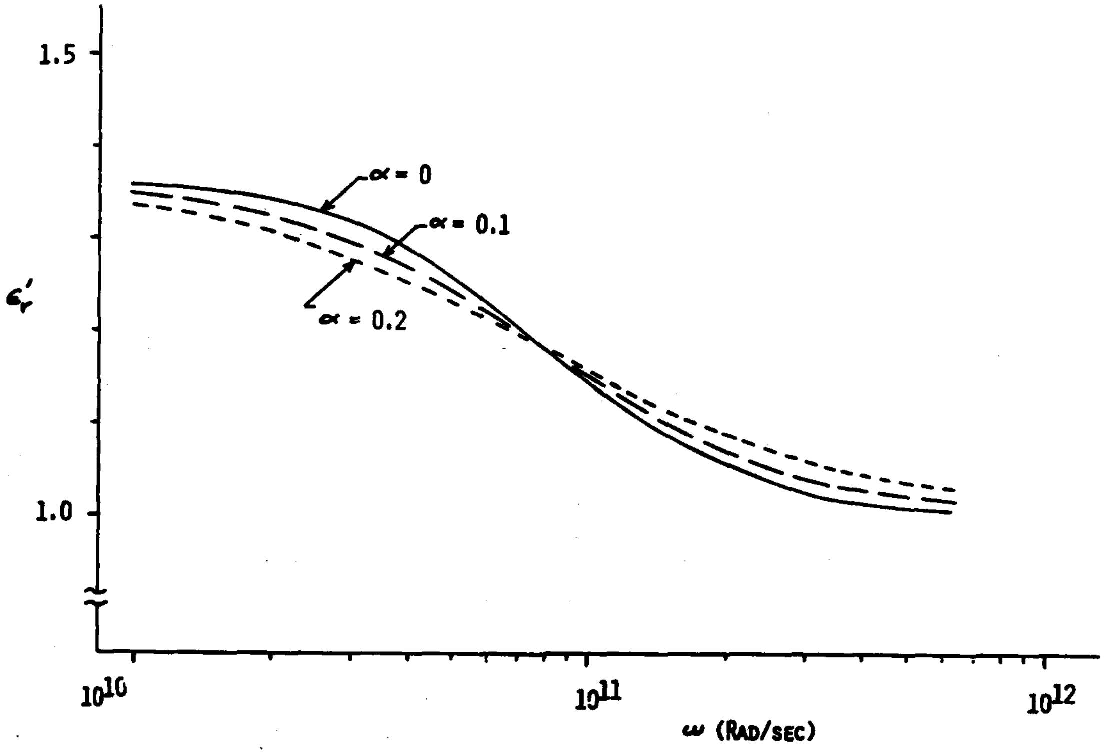


Figure 4. Real part of dielectric constant, Cole & Cole theory.

constant (real and imaginary parts shown in Figs. 4 and 5) should not greatly effect the reflectance of the RACO over a metal substrate.

VI. Summary

Theoretical calculations indicate that a new type of radar absorbing coating called a RACO could be an effective absorber of cm-mm wavelength radiation. A RACO is formed by compacting radar absorbing chaff (RAC) filaments and fixing them in a binder. Two types of RACO's are considered: regular and random. Regular RACO's are regular arrays of parallel filaments; random RACO's are random arrays of randomly oriented filaments.

In order to efficiently analyze RACO's the single filament RAC theory of Swinford (6) has been recast in terms of a relaxation formalism. This procedure has enabled us to write the dielectric constant for RACO's in terms of the Debye theory (7) of the dielectric constant from solid state physics. Having the Debye dielectric constant enables us to conveniently include dipole interactions for both regular and random RACO's.

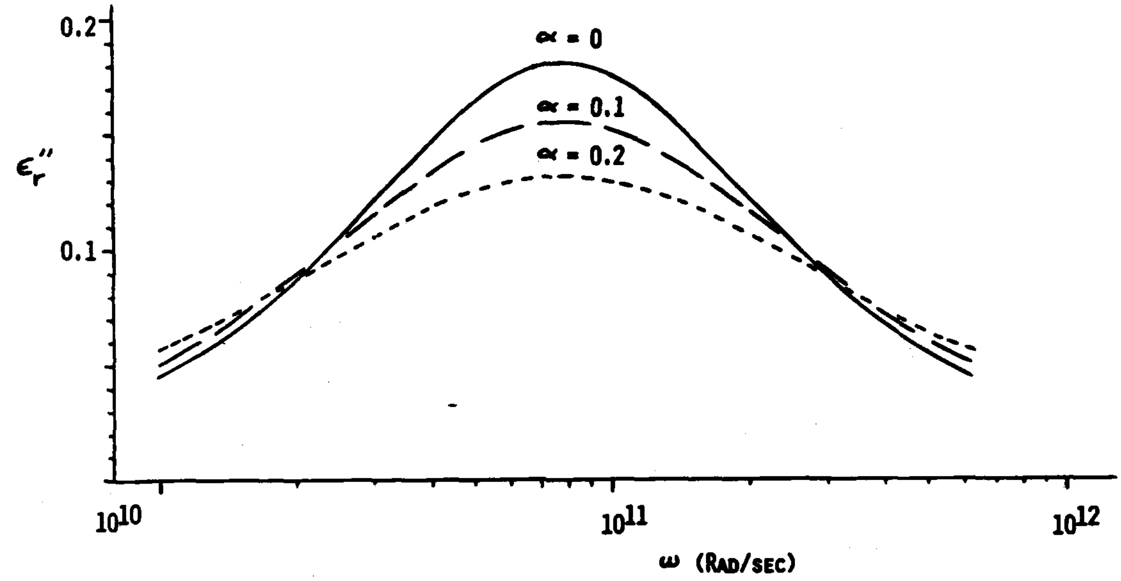


Figure 5. Imaginary part of dielectric constant, Cole & Cole theory.

The results of reflectance calculations for a representative regular RACO over a metal substrate are shown in Table I. Two versions of the Debye dielectric constant are used in the calculations. One version includes dipole interactions; the other version neglects all interactions (dipole, quadrupole, etc.). Reflectances computed using the two dielectric constants are in reasonable agreement. The reflectance is less than 3% at 30 GHz and falls much lower at higher frequencies. The RACO is a broadband absorber. The reflectance data are for normally incident EM waves. The RACO is 2 cm thick and the filaments are separated by one-third their length in the direction perpendicular to the filament axes.

At the close filament separations encountered in RACO's dipole and higher order interactions should be important. Why then do we include reflectance calculations that neglect interactions (dipole and higher order)? The answer lies in some engineering experiments (10) done in the 1950's on regular array artificial dielectrics. The results are shown in Fig. 2. The conclusion from these experiments is that for very close filament separations neglecting interactions is better than including dipole interactions. The agreement of the non-interactive theory with experiment is, of course, fortuitous. The interactive theory should be improved by the inclusion of interactions of higher order than dipole.

The results of reflectance calculations for a representative random RACO over a metal substrate are shown in Table II. Three theoretical expressions for the dielectric constant are used. Reflectances computed using all three expressions are in reasonable agreement. The reflectance is less than 3% at 30 GHz and drops to much lower values at higher frequencies. The reflectance data are for normally incident EM waves. The RACO is 1 cm thick and the filament average separation is about one-third the filament length.

Recent calculations indicate that even thinner RACO's are possible using a multi-layer coating. Thicknesses of one-half a centimeter or less appear possible while still giving broadband absorption.

The calculations described above have been done assuming a single relaxation time. In any real RACO there will be a distribution of filament lengths and thus a corresponding distribution of relaxation times. To handle this situation the Cole & Cole theory (9) of solid state physics has been adapted to the RACO problem. This theory demonstrates that in changing from a single relaxation time to a fairly broad distribution of relaxation times we produce little change in EM wave absorption by RACO's.

While our calculations indicate that RACO's are effective absorbers of cm-mm wavelength radar radiation we lack experimental confirmation for actual RACO material. Within the next few months laboratory work on RACO's should begin.

References

1. Severin, H., "Non-Reflecting Absorbers for Microwave Radiation," IRE Trans. Vol. AP-4, July 1956, pp 385-396.

2. Ryder, J.D., Networks, Lines, and Fields, Prentice-Hall, Inc., Englewood Cliffs, NJ, (1960), pp 462-465.

3. Skolnik, M.L., Introduction to Radar Systems, McGraw-Hill Book Company, Inc., New York, (1962), p 565.

4. Thomas, A.S., "Theory of Radar Absorber Material Design (U)," (A.S. Thomas, Inc.), Technical Report AFAL-TR-70-120, Sep 1970 (SECRET).

5. Mostrom, R.A., "Broadband Radar Absorbers (U)," (AVCO Systems Div., Wilmington, MA), Contract No. DAAB-07-70-C-0194, Aug 1972 (CONFIDENTIAL).

6. Swinford, H.W., "Electromagnetic Behavior of Radar Absorbing Chaff (RAC)," (Naval Weapons Center, China Lake, CA), Technical Note 354-43, June 1975.

7. Von Hippel, A.R., Dielectric Materials and Applications, The MIT Press, Cambridge, MA, (1954).

8. Onsager, L., "Electric Moments of Molecules in Liquids," J. Am. Chem. Soc., (1936), pp 1486-1493.

9. Cole, K.S., Cole, R.H., "Dispersion and Absorption in Dielectrics," J. Chem. Phys., (1941), pp 341-351.

10. Kharadly, M.M.Z., Jackson, W., "The Properties of Artificial Dielectrics Comprising Arrays of Conducting Elements," Proc. IEE (London), Vol. 100, Part III, July 1953, pp 199-212.

11. Kittel, C., Introduction to Solid State Physics, John Wiley & Sons, Inc., New York (1963).

12. Multilayer Normal Incidence Reflectance Program, programmed in Basic for Tektronix 4051 by A. Gauss, Jr., (to be published).

13. Blakemore, J.S., Solid State Physics, W.B. Saunders Company, Philadelphia, (1969), p 345.

14. Scott, W.T., The Physics of Electricity and Magnetism, John Wiley & Sons Inc., New York, (1959), pp 137-138.

HIGH ENERGY SULFURYL CHLORIDE BATTERIES (U)

SOL GILMAN, WILLIAM WADE, JR., and MICHAEL BINDER
US ARMY ELECTRONICS TECHNOLOGY AND DEVICES LABORATORY, ERADCOM
FORT MONMOUTH, NEW JERSEY 07703

INTRODUCTION

There is no practical alternative to the use of primary batteries to power man-portable electronic equipment for communications, surveillance, target acquisition, and Night Vision applications. For an increasing percentage of such newly-developed equipments, power and energy density requirements are so high that only the most energetic electrochemical couples can be considered for the purpose. Cells utilizing sulfuryl chloride as the cathode reactant are the latest and most energetic of the "liquid cathode cells" resulting from research initiated at the Electronics Technology and Devices Laboratory in the early 1970s.

Table I lists "full-cell" reactions and other electrochemical information for 5 different primary cell systems.

TABLE I - ELECTROCHEMICAL CELL REACTIONS

SYSTEM	CELL REACTION	Theoretical Potential E°(V)	Experimental OCV(V)	Theoretical Energy Density Wh/lb
Mg/MnO_2	$Mg + 2MnO_2 + H_2O \rightarrow Mn_2O_3 + Mg(OH)_2$	2.88	2.0	324
Li/SO_2	$2Li + 2SO_2 \rightarrow Li_2S_2O_4$	----	2.95	611*
$Li/SOCl_2$	$4Li + 2SOCl_2 \rightarrow 4LiCl + S + SO_2$	----	3.65	668*
Li/SO_2Cl_2	$2Li + SO_2Cl_2 \rightarrow 2LiCl + SO_2$	3.909	3.91	639
Ca/SO_2Cl_2	$Ca + SO_2Cl_2 \rightarrow CaCl_2 + SO_2$	3.818	3.30	584

*Based on open-circuit voltage.

The first system is the Army's present aqueous magnesium cell which follows conventional electrochemical practice in that the anode and cathode reactants are both solids and mechanically separated to interact only through an external electrical conductor. The other table entries are for "liquid cathode" cells which successfully violate the rule of separation of reactants. Although the anode and the cathode reactants are in direct contact (as the latter are liquids), their direct chemical reaction is kinetically hindered by a thin and protective layer of salt on the anode.

The relatively low theoretical energy density of the aqueous magnesium cell translates into a practical energy density of approximately 45 Wh/lb at moderate rates of discharge. In the nonaqueous systems it is possible to utilize much more reactive anodes and cathodes than in the aqueous magnesium system; therefore, higher electrode potentials, experimental voltages, and energy densities are feasible. For instance, the Li/SO_2 cell field-tested by our Laboratory actually provides an energy density of 100 Wh/lb at moderate rates of discharge. The $Li/SOCl_2$ cell is our first "oxychloride" system; it is already undergoing 6.3 stage development for target acquisition applications. It can deliver approximately 150 Wh/lb at moderate discharge rates. The Li/SO_2Cl_2 cell system, which will be discussed more extensively in this paper, has a theoretical energy density slightly lower than that of the $Li/SOCl_2$ cell. However, the higher terminal voltage and larger practical cathode capacities, which we have been able to obtain, allow us to project densities greater than 200 Wh/lb for a practical cell.

Several years ago, it was taken for granted that energy densities greater than 100 Wh/lb could be obtained only with cells utilizing lithium anodes. The last entry on Table I reveals that a Ca/SO_2Cl_2 cell system has the potentiality for performance in the same range as that of the lithium cells. Encouraging preliminary results on that system will also be reported here.

EXPERIMENTAL PROCEDURES

Preparation of Cells: Preparation of the electrolyte, electrodes, and assembly of the cell is discussed elsewhere (1). All steps and procedures involving the exposure of either the clean anodes or of the electrolyte were performed either in a glove box (argon atmosphere) or in a dry room. The preliminary "screening" of several different carbon black powders was performed on "uncompressed" electrodes made by a "standard" (1,2) technique, while the optimized United carbon electrodes were made using a cold-compression step in the fabrication process (1).

The cells were assembled using one 2 cm X 2.5 cm cathode sandwiched between two anodes of the same size with fiberglass filter papers providing mechanical saturation between the electrodes. A small lithium electrode was incorporated into the cell to serve as a reference electrode. All current densities are based on the total current divided by 5 cm^2. The cells were horizontally oriented on the bottom of an open teflon container with 5 cm^3 of electrolyte. The teflon container was, in turn, placed in a gas-tight glass outer container with electric feed-throughs.

Electrochemical Measurements: Discharge and polarization curves were measured at 22° ±2°C using a constant current power supply. The polarization curves were measured by applying pre-determined constant currents starting from 0.04 mA/cm^2 and allowing 3 minutes before recording each closed circuit potential. All potentials were measured against the lithium "reference" electrode.

Cathode Porosity Determinations: The percent porosity is defined as the percent of wet cathode volume available for absorption of SO_2Cl_2. The volume of SO_2Cl_2 was determined by weighing a cathode before and after immersion in SO_2Cl_2 and "blotting" on a glass surface. The wet volume of the cathode was determined by measuring its linear dimensions with calipers.

Determination of "True" Surface Areas: The surface area of carbon black powders and of complete electrodes was measured by the Brunauer, Emmett, Teller (BET) method using an Orr Model 2100 Surface Analyzer.

Electrolyte Conductivities: Conductivities were measured using an AC bridge and a thermostatted glass-stoppered cell with a cell constant of 0.200 cm^{-1}.

RESULTS AND DISCUSSION

Sulfuryl Chloride Electrolyte Solution: Of the salts with known solubility in SO_2Cl_2, $LiAlCl_4$ *has long-term chemical stability in solution,* has a solubility upwards of 2 moles/liter and does not undergo noticeable direct chemical reaction with lithium. Conductivities of a 1.5 molar solution of $LiAlCl_4$ in SO_2Cl_2 are plotted for a wide temperature range in Figure 1.

No precipitation of the solute was noted down to -54°C (the freezing period of pure SO_2Cl_2 is -54.1°C). At room temperature, the specific conductivity of approximately 0.01 Ω^{-1} cm^{-1} is of the same order as that encountered for the electrolyte used with other high energy battery systems. When using such electrolytes, internal voltage drops in cells are kept small by restricting separator and cathode thickness to less than a millimeter. The low conductivities encountered at -20°C or lower do affect performance significantly and improved electrolytes are being sought in our continuing in-house program.

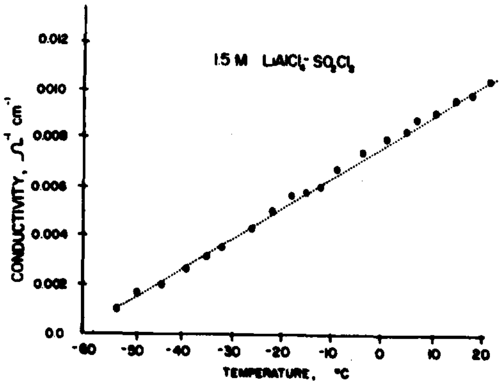


Figure 1. Conductivities of 1.5 M $LiAlCl_4$ - SO_2Cl_2.

Lithium Anode in Sulfuryl Chloride Cells: A lithium electrode with a freshly prepared (by mechanical abrasion) surface shows little voltage polarization for anodic current densities up to 40 mA/cm. On current reversal, dissolved Li^+ is reduced to metallic lithium and deposited on the surface of the lithium electrode in a highly dendritic form. Although precise thermodynamic measurements have not been performed, one may conclude that a fresh lithium electrode is close to reversible in a solution of $LiAlCl_4$ in SO_2Cl_2. The stability of the lithium electrode in SO_2Cl_2 is due to the formation of a self-limiting thin coating of LiCl, formed spontaneously according to the reaction in Table I. On long-term storage, that film apparently thickens and would introduce a "voltage delay" problem in a battery cell. This potential storage problem is presently under investigation by a contractor (3).

Teflon-Bonded Cathode in the Li/SO_2Cl_2 Cell: Electrodes for lithium thionyl chloride cells are normally formulated with Shawinigan carbon black and Teflon emulsion (normally 5-10% TFE in the dry electrode). Prior results for sulfuryl chloride cells, using similar electrodes, produced rather discouraging results (4). In the present work, we experimented with several different types of carbon black powders and with the process of electrode fabrication. Table II lists the carbons studied along with properties of the original powders and of completed electrodes fabricated using 16% TFE and a standard fabrication technique.

TABLE II. Teflon-Bonded Carbon Cathodes (16% TFE)

Type of Carbon	Derivation	Carbon loading* (g/cm^2)	Per gram of carbon powder	Per cm^2 of electrode geometric area*	% electrode porosity
Shawinigan—50% compressed	Decomposition of acetylene	0.0194	66	1.28	87
Darco-G60	Steam-activation of charcoal	0.0048	301	14.4	64
United XC-6310-4	Decomposition of oil	0.0146	1000	14.6	81
Columbia MR 1670	Decomposition of oil	0.0104	1200	12.5	75

* Based on (length x width) area of electrode, one side.

The resulting electrodes have "true" (BET) areas spanning more than an order of magnitude and electrode porosities ranging from 64 to 87%. Figure 2 presents polarization curves for such electrodes, with current densities based on the "superficial" area of the electrodes. From the figure, it is clear that very significant differences exist between the electrodes, with "United" carbon black affording the best initial voltage over most of the current range. The results are replotted in Figure 3 after basing the current density on the "true" (BET) rather than on the "superficial" surface areas. On that basis, it may be concluded that almost identical results are obtained for all of the carbons in the lower range of current densities where mass transport effects may be expected not to play a dominant role. This is consistent with the surface-chemical mechanistic sequence to be discussed below.

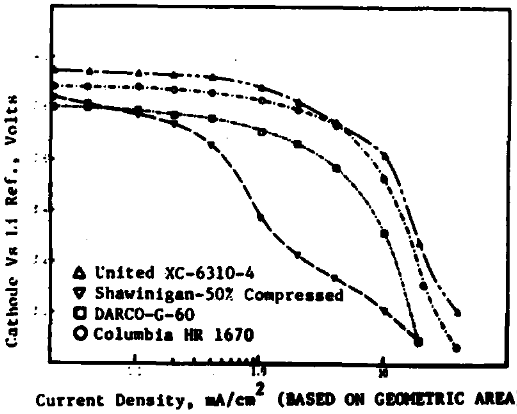


Figure 2. Polarization curves for Teflon-bonded carbon cathodes in lithium-sulfuryl chloride cells. Current densities based on "superficial" cathode areas.

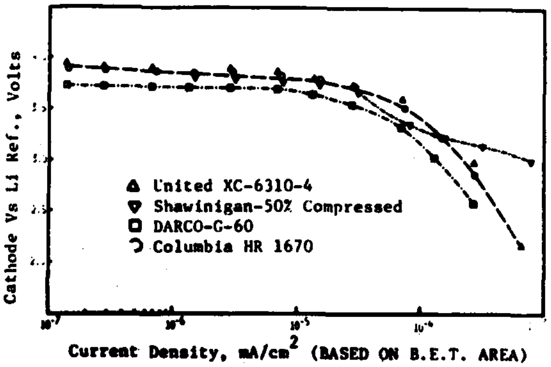


Figure 3. Polarization curves for Teflon-bonded carbon cathodes in lithium-sulfuryl chloride cells. Current densities based on "true" (BET) surface areas.

Figure 4 shows that the capacity of the test cathodes depends directly on the gross electrode porosity as measured by solvent absorption. The porosities are determined, to some extent, by the inherent microscopic "structure" of the carbon powders. Shawinigan black, for example, is noted for its highly developed acetylenic structure, carried over from the molecule from which it is originally derived.

From Figures 1-4 one can conclude that the "best" cathode would combine both high BET area and high porosity. Attempts were made to accomplish these conditions, simultaneously, by formulating cathodes with mixtures of Shawinigan black and the higher area carbons with unsatisfactory results. Success was achieved by varying the fabrication process in order to increase the porosity of United carbon electrodes. Figure 5 presents representative discharge curves for different formulations. The use of United Carbon and a "cold-compression" process (Trace 5, Figure 5) was adopted as standard for the balance of this work. The electrode was further optimized with respect to teflon content (at 11% TFE). A typical electrode had a porosity of 87% and a carbon loading of 0.024 gm/cm^2.

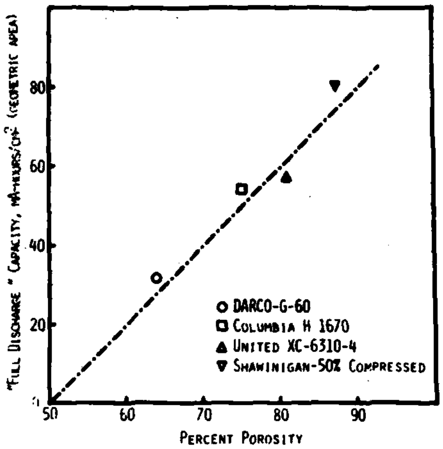


Figure 4. Capacity-porosity relationship for Teflon-bonded carbon cathodes (standard, "uncompressed" fabrication).

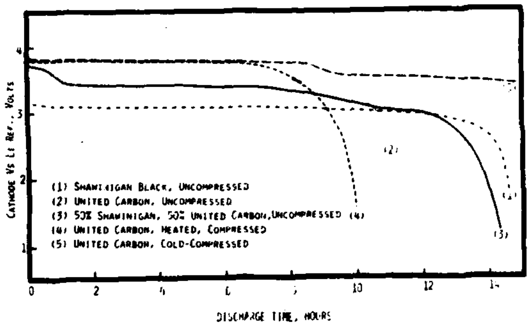


Figure 5. Discharge curves at current drains of 5 mA/cm^2 for differently-formulated cathodes in lithium-sulfuryl chloride cells.

Figures 6 and 7 compare performance of the cathodes of our Li/SO_2Cl_2 cell with that of the older Li/$SOCl_2$ system, which was previously the most energetic ambient temperature primary cell system known. Because anode polarizations are small, the voltages are essentially those of a complete practical cell. Although the performance illustrated in Figure 6 and 7 is excellent, there is still the possibility for significant improvements at high currents and low temperatures and that is one thrust of our continuing battery research program.

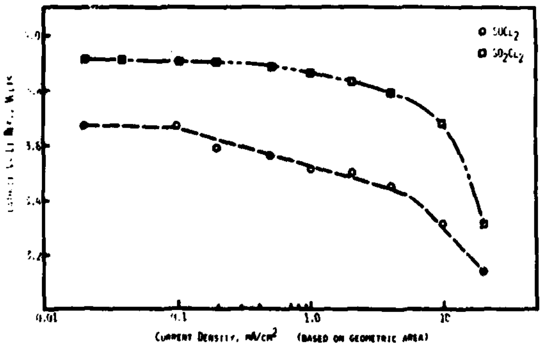


Figure 6. Comparison of cathodic polarization curves for lithium-sulfuryl chloride and lithium-thionyl chloride cells.

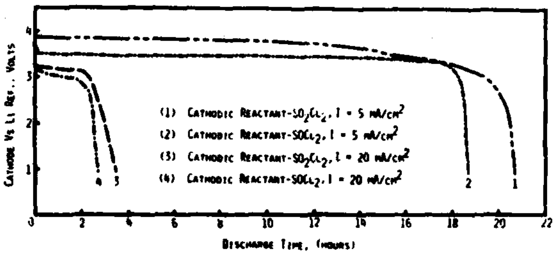


Figure 7. Comparison of cathodic discharge curves for lithium-sulfuryl chloride and lithium-thionyl chloride cells.

Electrode Discharge Products in the Li/$SOCl_2$ Cell: The overall discharge reaction listed for the Li/SO_2Cl_2 cell in Table I was deduced through x-ray diffraction, gravimetric chloride determinations and volumetric determinations of SO_2 gas release. First, it was determined that LiCl is, practically speaking, insoluble in the electrolyte solution employed here. Hence, discharged cathodes could be exhaustively washed with pure SO_2Cl_2 to eliminate entrained $LiAlCl_4$, without solubilizing LiCl formed as a product in the cathode. Diffraction patterns for cathodes which were carefully washed, dried and crushed revealed only cubic LiCl. Two cells made with United carbon cathodes were discharged at a current density of 5 mA/cm^2 to a cutoff voltage of 2 Volts. The cathodes were washed with SO_2Cl_2, dried and quantitatively extracted with water. Aliquots of the aqueous solutions were potentiometrically titrated for Cl^- using a standardized $AgNO_3$ solution. A small correction for occluded $AlCl_4^-$ was made by determining Al^{+3} colorimetrically and subtracting four times that number of equivalents. The corrected equivalents of Cl^- were found equivalent to the coulombs of electricity passed, to within a 2% tolerance. This establishes that LiCl is the only ionic product of cell discharge and also that the LiCl is quantitatively precipitated in the pores of the cathode.

To evaluate non-ionic discharge products, it was first determined quantitatively that gaseous SO_2 is released during cell discharge. A cell was then assembled with electrolyte which was pre-saturated with SO_2. The cell was attached to a gas buret (using fluorocarbon oil as the displacement liquid), and discharged at 5 mA/cm^2 while monitoring the volume of gas released. In duplicate experiments, it was established that one mole of gas was produced for every two equivalents of electricity passed.

The overall reaction listed in Table I is the simplest one consistent with the analytical results reported above. The following individual electrode reactions are consistent with the overall reaction:

$$\text{anode:} \quad 2[Li \rightarrow Li^+ + e^-] \tag{1a}$$

$$\text{cathode:} \quad 2Li^+ + SO_2Cl_2 + 2e^- \rightarrow 2LiCl + SO_2 \tag{1b}$$

$$\text{overall:} \quad 2Li + SO_2Cl_2 \rightarrow 2LiCl + SO_2 \tag{2}$$

$$E^\circ = 3.909 \text{ V } (30^\circ C)$$

As already discussed above, the capacity of the cathode is proportional to its porosity. As is apparent from these findings, the porosity is required in order to accommodate the LiCl precipitate while allowing continued good transport of the cathode reactant and of conducting ions.

The analytical results reported above contradict earlier reports (5) that Li_2SO_4 and molecular sulfur are major products of cell discharge. It is felt that the earlier results may have resulted from a failure to avoid water contamination early in the analytical procedure. The present results are particularly significant for their relevance to battery safety, inasmuch as molten sulfur can react vigorously with lithium under conditions of cell malfunction. This is potentially a very significant advantage of the Li/SO_2Cl_2 over the older Li/$SOCl_2$ system (see Table I).

Cathode Mechanism in the Li/SO_2Cl_2 Cell: Sulfuryl chloride is known (6) to dissociate into SO_2Cl_2 according to the following reaction:

$$SO_2Cl_2 \rightleftharpoons SO_2 + Cl_2 \qquad (3)$$

Both gases are appreciably soluble in our electrolyte (1.09 and 0.62 molal for SO_2 and Cl_2, respectively, at 24°C as determined in this laboratory). Furthermore, as determined by W. Behl (7), of this laboratory, molecular Cl_2 is more reactive than undissociated SO_2Cl_2. Figure 8 shows that saturation of the electrolyte with either SO_2 or Cl_2 produces a noticeable shift in the cathode polarization curve. The particularly marked effect of SO_2 implies that the latter is acting by removing Cl_2 as cathode reactant and shifting the equilibrium of reaction (3) to the left. A complete cathode mechanism, consistent with our observations, would involve the production of Cl_2 by reaction (3) followed by the cathodic reduction of Cl_2:

$$Cl_2 + 2e^- + 2Li^+ \longrightarrow 2LiCl \qquad (4)$$

Carbon is a known catalyst for reaction (3) and that would explain the importance of high electrode surface area to good electrode performance. Quantitative information on reaction (3), in the condensed phase, is desired for confirmation of the proposed cathode mechanism and as a basis for future cathode improvements. Preliminary studies have been performed using ultraviolet spectrophotometry. Figure 9 shows that "aged" sulfuryl chloride absorbs light with wavelengths less than 460 mμ (the liquid appears yellow). The absorbance is diminished by freshly-saturating with argon, and further diminished by saturating with SO_2 (the latter does not absorb light in this region of the spectrum). The absorbance is intensified by adding Cl_2. Comparison of the results suggests that the "aged" solution has undergone decomposition. Estimates of the rate of decomposition can be made by recording the absorbance (e.g., at 400 mμ) vs time relationship for freshly degassed samples. Preliminary results reveal that, at 30°C, the homogeneous rate for the "neat" solvent is equivalent to less than 0.2 mA/cm² (exact rates depend on solvent purity) and, hence, would not support good cathode performance at high current densities. In the

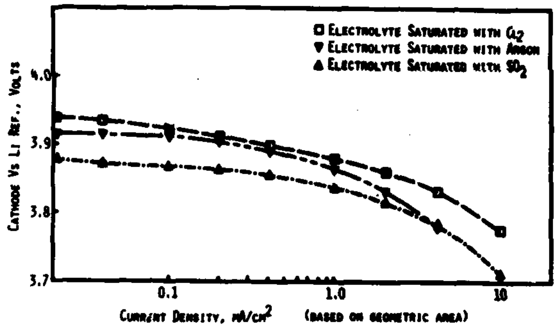


Figure 8. Cathodic polarization curves of lithium-sulfuryl chloride cells containing gas-saturated electrolyte.

presence of carbon, however, the estimated heterogeneous rate of decomposition is sufficiently large to account for the full range of current densities covered in Figure 8.

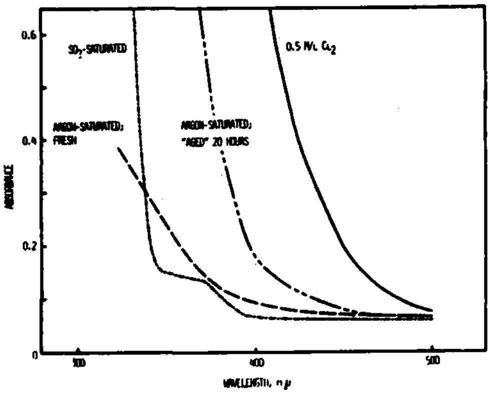


Figure 9. Ultraviolet spectra of sulfuryl chloride samples.

Performance of Electrodes in the Ca/SO_2Cl_2 Cell: The overall discharge reaction listed for this cell in Table I is based on analogy to the Li/SO_2Cl_2 cell. For that assumed reaction, the thermodynamic cell potential is only 91 millivolts less than that for the analogous lithium system. From Figure 10, it can be seen that experimental cell voltages are always at least ½ Volt less than the thermodynamic value of 3.818 Volts. Furthermore, most of the low current density polarization occurs at the calcium anode, which might otherwise be expected to assume a potential approximately 90 millivolts positive to the lithium reference electrode. However, it is highly encouraging that the calcium anode shows little tendency to undergo additional polarization at high current densities (i.e., it holds promise for good "high-rate performance). The cell polarization at high current densities is almost entirely attributable to the cathode, similar to the situation for the Li/SO_2Cl_2 cell.

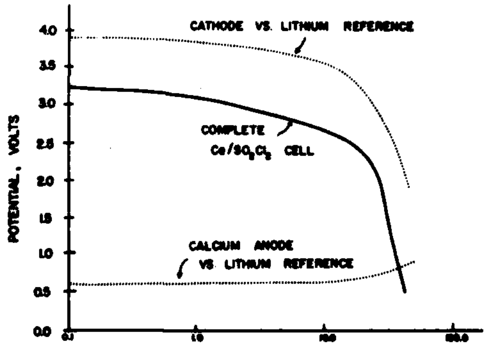


Figure 10. Polarization curves for a calcium-sulfuryl chloride cell.

Figure 11 compares cell polarizations for our Ca/SO_2Cl_2 cell with those for our Li/SO_2Cl_2 cell and with a commercial Li/SO_2 cell of the type presently being used in Army radio sets. From the figure, it is clear that the new calcium cell, while still inferior to our best lithium technology, has the potentiality for displacing the Army's present lithium-sulfur dioxide cell for at least some applications. This may apply particularly where the greater chemical stability of calcium outweighs the superior electrochemical performance of lithium.

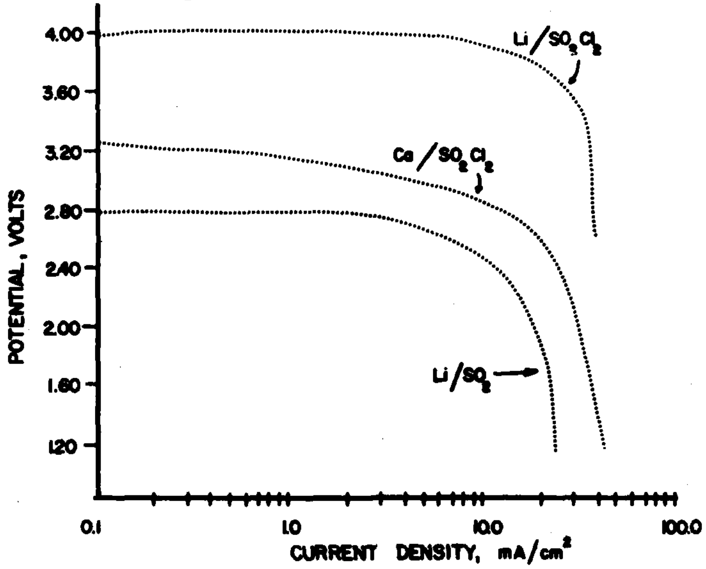
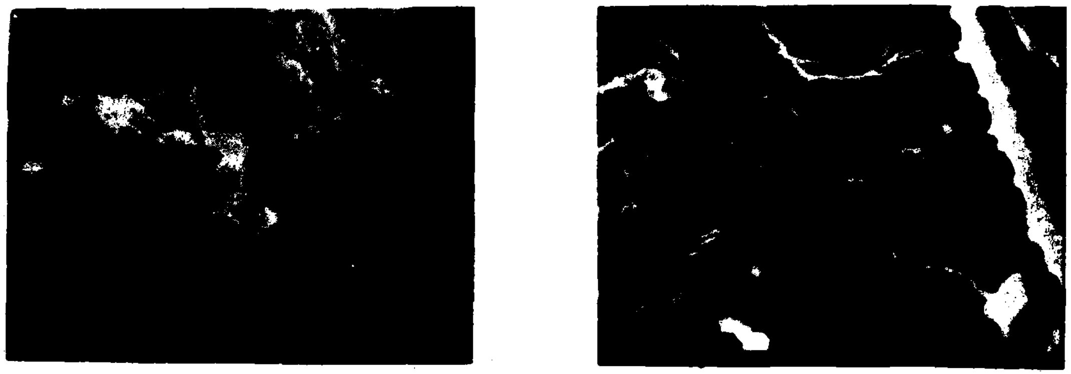


Figure 11. Polarization curves for three different high energy cells.

Although the polarization characteristics of the electrodes in a calcium-sulfuryl chloride cell are highly encouraging, we are presently faced with the serious problem of reduced cathode capacity (i.e., life) in any electrolyte containing dissolved calcium ions. Presently, cathode life is one-half or less of that for the analogous lithium cell. Micrographs of spent cathodes from the two types of cells (Figure 12) provide a clue to the origin of the problem. In the lithium cell, the LiCl which forms in the cathode during discharge, is deposited as multiple discrete clusters of cubic crystals which allow relatively easy passage of ions, reactants and products from the liquid electrolyte phase. By contrast, the solid product phase in the cathode from the calcium cell appears continuous and "glassy" and would be expected to interfere with transport and conduction processes in the electrolyte phase. Studies are currently in progress to evaluate and overcome this problem.



a. Lithium anode

b. Calcium anode

Figure 12. Scanning electron micrographs for fully discharged cathodes in SO_2Cl_2 (3200 X magnification).

Finally it may be noted that on thermodynamic grounds, approximately 500 millivolts of terminal voltage can be "recaptured" if the calcium electrode can be made to operate more reversibly. The present performance, similar to that observed in the lower-potential Ca/$SOCl_2$ system (8,9) can be attributed to "film polarization" in the protective coating (probably $CaCl_2$, according to the reaction in Table I) formed by limited spontaneous reaction of the anode with the solvent. Improved terminal voltage can be expected if it is discovered how to minimize the latter coating while still retaining its protective properties.

REFERENCES

1. S. Gilman and W. Wade, Jr., Electrochem Soc., 127, 1427 (1980).
2. W. K. Behl, J. A. Christopulos, M. Ramirez and S. Gilman, J. Electrochem Soc., 120, 1619 (1973).
3. F. Marakar, "High Rate Lithium-Sulfuryl Chloride Battery Technology," First Quarterly Report, Contract DAAK20-81-C-0420 (ERADCOM), Gould Labs, March 1982.
4. J. J. Auborn and N. Marincic, "Power Sources 5," D. H. Collins, Editor, p. 683, Academic Press, London (1975).
5. J. J. Auborn, R. D. Bezman, K. W. French, A. Heller and S. F. Lieberman in Proceedings of 26th Power Sources Symposium, p. 45 (1974).
6. "Encyclopedia of Chemical Technology," Vol. 14, K. Othmen, Editor, p. 45 (1969).
7. W. K. Behl, J. Electrochem Soc., 127, 1444 (1980).
8. R. L. Higgins, Proceedings of the 29th Power Sources Conference, p. 147 (1980).
9. R. J. Staniewicz, J. Electrochem Soc., 127, 782 (1980).

THERMAL CAMOUFLAGE OF FIXED INSTALLATIONS: PROJECT TIREX (U)

*CURTIS L. GLADEN, MR.
LEWIS E. LINK, DR.
USAE WATERWAYS EXPERIMENT STATION
VICKSBURG, MS 39180

I. Introduction

1. Camouflage of fixed installations is a special problem because:

   a. The facilities are static and their location and configuration are often well known.

   b. Individual structures may be quite large and groups of structures can cover an extensive ground area.

   c. Few fixed installations have been constructed or sited properly for the expedient addition of camouflage measures; similarly, new facilities seldom include camouflage in their design criteria.

   d. Very little attention has been focused on development of materials and concepts for camouflage of fixed installations since World War II.

Increased adversary air strength and advancements in target acquisition and terminal homing systems have placed a new premium on camouflage and deception at fixed installations. The Corps of Engineers is addressing this requirement in its fixed-installation camouflage research program.

2. The Corps research has considered a number of possible approaches for camouflaging a fixed installation. Any given approach must consider both camouflage technique and the scope of camouflage application. Technique concerns alternative camouflage methods such as tone-down, shape disruption, and decoys. Scope of application defines the degree to which camouflage is applied such as the number of structures or the proportion of the entire installation on which camouflage is used.

3. The most extensive application of camouflage would include an entire installation. This could involve a disguise of the presence of the installation (primarily to defeat surveillance systems) or to a lesser degree an overall reduction in the conspicuousness of the installation to defeat target-acquisition systems. A less extensive approach would be to apply camouflage only to the critical elements of the installation and those features in proximity to the critical elements that might be used to estimate their locations. Critical elements in this sense are those necessary for the survival of resources required to execute the operational mission of the installation or to retaliate after an attack. The latter approach lends itself specifically to the defeat of target acquisition and terminal homing devices used with manned fighter-bomber aircraft.

4. Manned fighter-bomber aircraft approaching at low altitudes and high rates of speed require a certain amount of time to locate specific targets and lock-on with a weapon system. This can be translated to distance; that is, the pilot, visually or with a target-acquisition aid, must detect and identify a specific target at a sufficient range from that target to allow lock-on and accurate firing or weapon release. It is against this threat that camouflage of fixed installations can provide the most cost effective counter and increase the survivability of an installation's critical resources. The Corps' ongoing research emphasizes the development of thermal camouflage measures against this threat to complement existing techniques used in the visual band.

5. Camouflaging structures to reduce their thermal contrast to the surrounding terrain is particularly difficult because of the significant and at times rapid variations in both target and background thermal infrared (IR) signatures. Targets can be warmer than the background during one portion of the day and cooler a few hours later. A change in the weather can reduce target-background contrast, eliminate it, or increase it within hours or a fraction of an hour. Since many target-acquisition systems utilize imaging sensors, the geometries (shape and texture) of the target and background are also important considerations. Terrain areas surrounding structures can vary considerably in their pattern and texture as perceived by thermal IR sensors. An area having variations in the percentage of grass cover and adjacent patches of trees may appear very uniform in tone with subtle boundaries at one time and heterogeneous with significant differences in tone and texture both between and within the areas of grass and trees at another time. Effective camouflage of a structure within such an area may require that the pattern and texture of the camouflage vary in phase with that of the background.

6. The growing threat posed by thermal IR imaging systems and heat seeking munitions and the almost complete dearth of information on thermal camouflage of fixed installations are the impetus for the Corps' research emphasis in this area. The interest in developing thermal camouflage

technology is shared by many of our NATO allies and formed the basis for the first major thermal camouflage field trial focused on fixed installations. This paper provides an overview of the field trial called Thermal Infrared Experiment (TIREX) and presents some preliminary results.

II. Overview of Project TIREX

A. Organization and Approach

7. In February 1979, members of NATO Working Group D of the Special Group of Experts on Concealment, Camouflage, and Deception (NATO AC/243 (CCD) WG(D)) agreed to cooperate in Project TIREX. One objective of the trial was to evaluate the potential of a variety of materials and material application concepts to reduce the vulnerability of critical elements at fixed installations to air attack by systems using thermal IR sensors.

8. Project TIREX participants included Denmark, the Netherlands, the United Kingdom, the United States, and West Germany. The U.S. Army Engineer Waterways Experiment Station (WES), Vicksburg, Mississippi, U.S.A., had responsibility for overall coordination of project TIREX. Participation in the program involved contributing samples of thermal camouflage material for testing, supplying measurement teams for gathering field data, and supplying scientific personnel to analyze the data collected. The U.S. Air Force Base at Zweibruecken, West Germany, was selected as the site for the field trial. Because of the size of the installation and the diversity of its structures, no effort was made to examine and treat the whole installation or even significant parts of it. Instead, two general target categories were decided upon and several targets in each category were examined. The two categories considered were hot targets and cold targets. Hot targets were defined as buildings with internal heating and minimal insulation. Cold targets were unheated structures such as aircraft shelters and taxiways.

9. A two-pronged method of problem attack using a combination of computer modeling and field experiments was selected. The field experiment portion of the study was conducted in two phases. Phase 1 (Aug-Oct 79) involved the gathering of baseline temperature and thermal imagery data at the field site. Phase 2 (Sep-Nov 80) involved a side-by-side field comparison test of various potential thermal camouflage materials. Computer modeling will be used to extend the field test results to other climatic conditions, to other camouflage materials, and to other camouflage configurations. The final products will include initial guidance for the design and application of thermal camouflage for critical elements of fixed installations.

B. Test Structures

10. A variety of camouflage measures, some currently in use and others under development, were applied in different manners to four structures at the air base. The structures involved were a large metal building, an aircraft shelter, a hardstand, and a hardened concrete building. The hardened concrete building, the aircraft shelter, and the hardstand were considered cold targets. The large metal (corrugated) building was considered a hot target because of its internal heating and lack of insulation. These structures were not camouflaged completely. Instead, a number of different materials or combinations of materials were applied to each structure so that the performance of the various measures could be observed with respect to one another. In each case, a sizeable area of the structure was left uncamouflaged to provide a reference. The following paragraphs provide a brief description of the camouflage measures used on each structure.

11. Figure 1 is a photo showing the deployment of camouflage measures on the corrugated metal building east wall. Additional camouflage

Figure 1. Photo of corrugated metal building from the east

measures were applied to the roof and to the roof of the bunker in the foreground. The netting stretched over the chimney is the West German Far Infrared (FIR) net system. Next to it on the wall is the Dutch low emissivity net over a layer of bubble foil insulation. The bubble foil served as wall insulation which was attached to the outside of the wall rather than the inside. Next on the wall is a fairly large darker colored area. This area was painted with a U.S. camouflage paint having a high near-infrared reflectance in order to reflect away a portion of the sun's energy. The right half of this area also had fiberglass insulation installed inside the wall. Next on the wall is a somewhat lighter colored area. This was the untreated reference area. The dark material on the right corner of the building is the West German OGUS mat. The OGUS mat has a foam insulator sandwiched between an olive drab aluminized layer and an impregnated olive drab cloth layer. In the right center of the picture is a sloped surface. This is the east facing bunker roof which was mentioned previously. The left darker colored side is the untreated asphalt reference surface. On the right side is an area painted with U.K. near infrared reflecting (NIRR) paint.

12. The hardstand is a concrete pad typical of the hardstands used for aircraft parking and for equipment storage at Zweibruecken. Camouflage materials were applied to the entrance of the hardstand as shown in Figure 2. In the foreground, the darker area covering the entire bottom

Figure 2. Photo (from west) of hardstand with camouflage measures tested

of the picture is the Danish texture mat, a synthetic carpetlike material. The light-colored area in the left center of the picture is covered with yellow paint. The yellow paint is used for taxiway marking, centerline stripes, and so on. It is highly reflecting in the solar band thus preventing heating of a surface on which it is placed. Next to the yellow paint is an area with U.S. green camouflage paint and to the right of it is an area with tan U.S. camouflage paint. Beyond the white square the dark area at the left center of the picture is another of 3M Company black paint. This was painted on the hardstand to serve as a hot reference surface with known surface properties (emissivity and absorptivity) for the aerial thermal imagery. The dark area to the right of the 3M paint is covered with Dutch low emissivity paint. The area beyond the barrel in the left center of the picture is the untreated concrete which served as the reference surface.

13. Camouflage materials were applied over the entire roof arch of the aircraft shelter so that the north wall, the roof, and the south wall were covered by strips of material. Figure 3 is a photo of the site as seen from the south side. Starting from the left, the first camouflage measure which is seen is the Dutch low emissivity paint. It shows in the photo as a lighter colored area. Instead of simply applying a strip of the paint, an effort was made to also apply a pattern with the paint. This shows as the bulge at the base of the shelter. Moving toward the

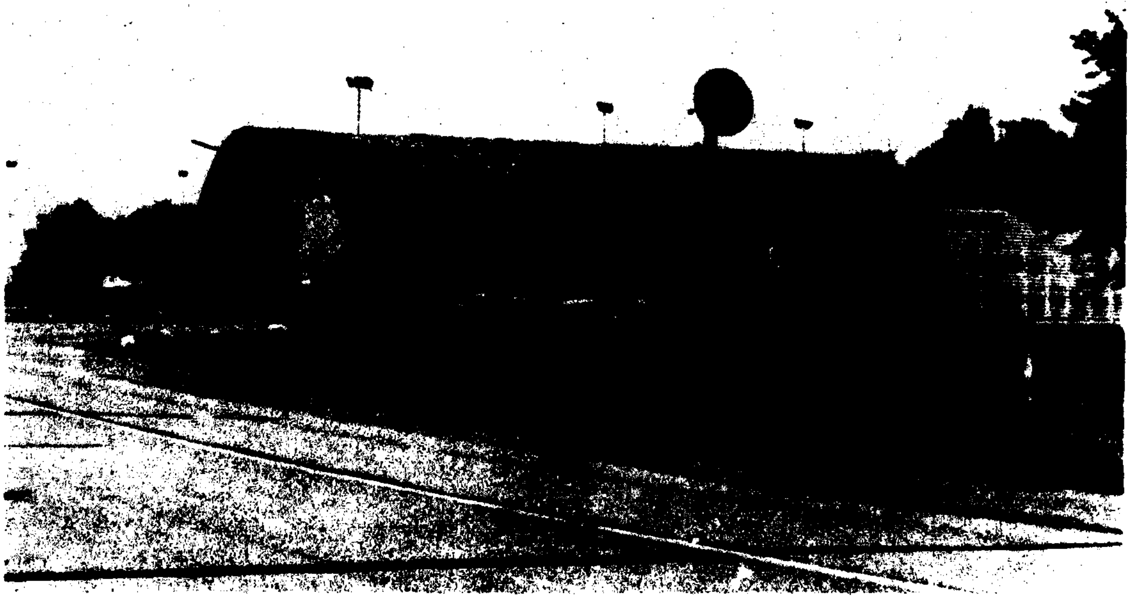


Figure 3. Photo of camouflage materials on aircraft shelter

right, the next area seen is the untreated reference area. The next camouflage measure was a combination of an old U.S. jute camouflage net with water trickling over it. The darker area to the right of the jute net is an area of untreated concrete shelter with water trickling over it. The next material with the white rope crosses on it is the Danish texture mat. The remainder of the shelter to the right was covered by a U.K. camouflage net that was sparsely garnished. This material added some elements of texture to the surface. The left half of this netted area was a combination measure. Before the netting was put in place, a U.K. low emissivity foil was applied on the surface.

14. The hardened concrete structure was chosen as representative of an aboveground hardened facility. Camouflage materials were applied to the roof and to the west wall. Some of the roof materials overlapped the east wall somewhat. Figure 4 is a photo of the site which shows the west wall. On the left side of the wall is the sparsely garnished U.K. camouflage net. This measure was staked out about 5 m from the wall and formed a sloping structure. The right half of this structure was cooled by water sprinkling. The next camouflage measure was an effort at cooling the wall by semiforced convection. This shows as the area of 3 m x 3 m squares. A frame was attached to the wall and slats were inserted into the frame in a way similar to a venetian blind except that the angle of the slats was fixed. Some of the slats were inclined downward and some were inclined upward. Two types of slats were used. The three darker squares had green painted wooden slats and the six lighter squares had asbestos slats.

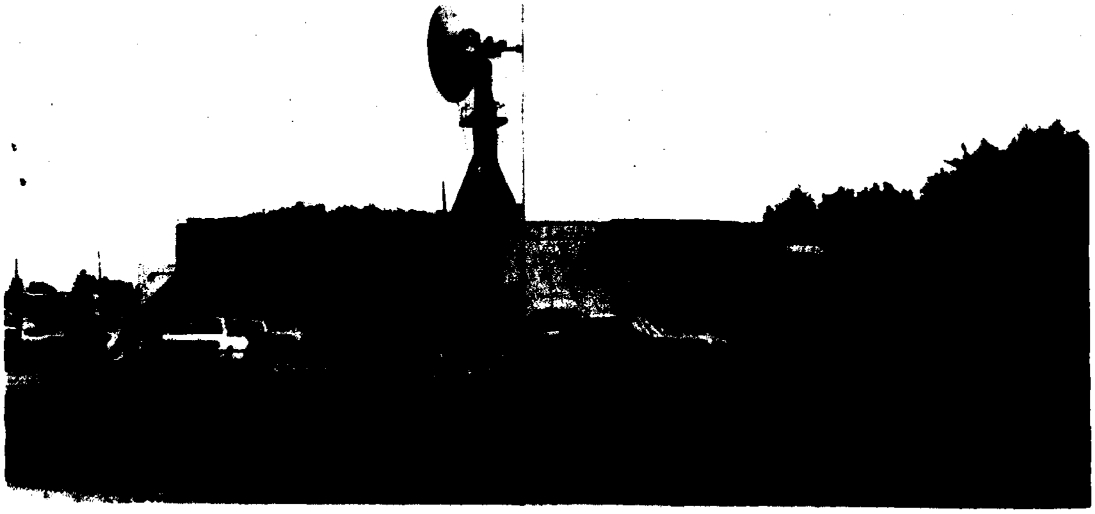


Figure 4. Application of camouflage materials on hardened concrete building

15. The light-colored area to the right of the slats was the untreated wall reference area. Next to it is a combination measure. The top half of the wall was painted with Dutch low emissivity paint. Then half of this area was covered by a Dutch camouflage net which was suspended vertically at a distance of about 1 m from the wall. The lower half of the wall was covered by Danish texture mat which was attached directly to the wall. Next to this measure is another combination measure. Beneath the Dutch camouflage netting seen in the photo is a U.S. furnished mat made up of space blankets, a commercially available thermal blanket. The space blankets were hung with the shiny side facing out in order to form a low emissivity surface. The last measure on the west wall was also a combination measure. A U.S. thermal blanket was draped down next to the wall and then a U.S. camouflage net was stretched out over it away from the wall. This measure appears in the photo immediately to the right of the white wall column next to the trees.

C. Ground Measurements

16. The ground measurements were made to characterize site conditions, thermal characteristics of the camouflaged and uncamouflaged structures, and background features over diurnal cycles and in different weather conditions. Four types of ground data were obtained: weather, feature temperature, feature reflectance, and imagery.

17. The weather data consisted of measurements of solar insolation, wind speed, wind direction, relative humidity, and air temperature. Measurements were recorded every 30 minutes with each value representing an average value for the previous 30 minutes. Measurements were recorded on magnetic tape cassettes for subsequent analysis.

18. Contact surface temperature measurements were made for most camouflage measures, uncamouflaged reference areas, and adjacent background areas using thermistors attached to the material surfaces. The thermistors were attached to data logger and recorder systems similar to the one used for the weather data. Temperature values were recorded every 30 minutes for the duration of the experiment.

19. Radiometric temperature values were obtained manually by use of an AGA Thermopoint TPT 80 Radiation Thermometer, a Wahl Heatspy Radiation Thermometer, and a Barnes Engineering Instatherm hand-held measurement device. A digitron Model 1751 Contact Thermometer was used to supplement these devices. Radiation temperatures were obtained over selected diurnal cycles spaced over the extent of the field data collection. Periodic measurements were made of the surface radiometric temperatures of the individual camouflage materials, uncamouflaged reference areas, and background features adjacent to the structures camouflaged.

20. Optical reflectance measurements for the individual camouflage materials (in situ), reference areas, and background features were made periodically by the various scientific teams. Additional material properties were determined through laboratory analyses by IABG Corporation, West Germany. The optical reflectance measurements were made with an integrating quantum radiometer/photometer providing hemispherical reflectance for solar (.4 to 1.2 μm), near infrared (0.67 to 0.85 μm), and visual (standard response of the human eye) bands.

21. Thermal IR imagery was obtained from preestablished ground stations by two West Germany teams (FfO and IABG) and teams from the Netherlands, Denmark, and the United States. The thermal IR imagery was acquired with modified AGA Thermovision Systems (Denmark and West Germany), Inframetric (United States), and a privately designed system used by the Netherlands team.

22. The image information was recorded on magnetic tape for most of the systems with the exception of the IABG system, which used an analog to digital converter to generate digital records of the imagery. All measurements were made in the 8- to 14-μm spectral band with the exception of those made by the Netherlands, which were made in both the 3- to 5-μm and 8- to 14-μm spectral bands. The various scientific teams acquired imagery over selected diurnal cycles to provide relatively uniform coverage over the full period of the field data acquisition. Since all teams acquired images from the same positions and at approximately the same times during the diurnal cycle, the imagery collected form a compatible data set that can be used to examine the performance of the respective camouflage measures over diurnal cycles and for a variety of weather conditions that span the 15 Sep to 10 Nov time period.

D. Aerial Data

23. Aerial overflights were made to acquire both visual and thermal IR imagery. The visual imagery consisted of conventional panchromatic aerial photography. Two types of thermal IR imagery were obtained: vertical line scan imagery and forward-looking infrared (FLIR) imagery. Vertical line scan imagery was acquired by U.S. Air Force tactical reconnaissance aircraft. FLIR imagery was acquired by the West German FfO team using a Honeywell mini-FLIR system. The team acquired simulated FLIR imagery with the Inframetrics imaging system mounted in a helicopter. A limited amount of FLIR imagery was acquired in 1979 prior to the application of camouflage materials by a U.K. aircraft. The FLIR imagery was obtained at different times of day and along preselected flight paths.

III. Data Analysis

A. Analysis Approach

24. The major objectives of the data analysis were to (1) evaluate the performance of the various camouflage materials for contrast reduction, both visual and thermal IR; and (2) evaluate the potential of the various camouflage measures (materials applied in specific ways) to reduce the detectability of specific types of structures for a variety of time and weather conditions. The analysis of contrast reduction potential is being accomplished primarily with the ground radiometric temperature, reflectance, and contact temperature data. The analysis of detectability reduction potential is being accomplished primarily with imagery data supplemented by modeling and simulation techniques.

25. Analyses of the data collected during the 1979 and 1980 field trials has been ongoing by scientific teams from each nation. The analysis of contrast and reduction potential is largely completed, while the analysis of detectability reduction potential is ongoing. Because of the sheer volume of data collected by the various scientific teams, it is practical only to present a few example or representative analyses in this paper. The analyses presented are those accomplished by WES personnel on the U.S. ground based radiometric temperatures, reflectance, and contact temperature data for the aircraft shelter.

B. Analyses of Radiation Temperature Data

26. For each data point (set of instrument readings), an average temperature of the material being examined was computed by adding the temperature readings from the various instruments together and dividing by the number of instruments used. Next, the absolute value of the temperature difference between the camouflage or reference material and its appropriate background material (grass or trees) was calculated. Two measures of camouflage effectiveness were calculated based on these absolute value calculations. The first of these measures was a daily sum of the absolute values for each camouflage measure and each reference surface. This provides a measure of the total daily temperature differences between the camouflage or reference material and its background. The second of these measures was the daily sum for each camouflage material divided by the daily sum for its appropriate reference material. This quotient was subtracted from one. This measure compares the camouflage and reference materials. If the result of this calculation is one, then the camouflage is perfect. However, if the result of this calculation is less than zero, then the camouflage material is less effective than the untreated reference surface. Table 1 presents the results from the aircraft shelter for three data collection days. The last columns in the table provide overall measures for all three days. Based on these data the wet concrete and wet

Table 1

Radiation Temperature Measures of Effectiveness (MOE) for

Camouflage Materials for Aircraft Shelter

Camouflage or Reference Material	1 Oct 80		25 Oct 80		31 Oct 80		Summary	
	MOE#1	MOE#2	MOE#1	MOE#2	MOE#1	MOE#2	MOE#1	MOE#2
North Wall								
Net (U.K.)	4.1	0.2	0.8	-0.3	3.7	-0.2	8.6	0
Screen and Net (U.K.)	5.9	-0.2	2.5	-3.2	14.6	-3.9	23.0	-1.7
Mat (Denmark)	3.5	0.3	1.2	-1.0	4.6	-0.5	9.3	-0.1
Wet Concrete	3.6	0.3	0.6	0	1.4	0.5	5.6	0.3
Wet Jute (U.S.)	3.1	0.4	1.2	-1.0	1.8	0.4	6.1	0.3
Reference Wall	5.0	0	0.6	0	3.0	0	8.6	0
LE* Paint (Netherlands)	6.4	-0.3	1.6	-1.7	1.9	0.4	9.9	-0.2
Reference West Hardstand	10.4	0	1.5	0	5.6	0	17.5	0
West Hardstand LE Paint	16.6	-0.6	3.3	-1.2	25.9	-3.6	45.8	-1.6
South Wall								
LE Paint (Netherlands)	11.1	0.6	2.8	-3.7	8.0	0.8	21.9	0.7
Reference Wall	31.3	0	0.6	0	36.9	0	68.8	0
Wet Jute (U.S.)	14.7	0.5	1.3	-1.2	19.5	0.5	35.5	0.5
Wet Concrete	15.8	0.5	0.7	-0.2	23.1	0.4	39.6	0.4
Mat (Denmark)	41.9	-0.3	2.1	-2.5	41.6	-0.1	85.6	-0.2
Screen and Net (U.K.)	16.7	0.5	4.7	-6.8	20.7	0.4	42.1	0.4
Net (U.K.)	7.6	0.8	1.3	-1.2	8.8	0.8	17.7	0.7

Note: $MOE\#1 = \Sigma|C-B|$ and $MOE\#2 = 1 - (\Sigma|C-B|)/(\Sigma|R-B|)$

where C = temperature of camouflage material
B = temperature of background (grass or trees)
R = temperature of untreated reference surfaces

*LE = low emissivity

jute (old U.S. camouflage net) were the most effective camouflage measures deployed on the north wall. On the south wall, the British camouflage net was the most effective camouflage measure, followed closely by the Dutch low emissivity paint. The result for the British net may be misleading because it was shaded from direct sunlight during portions of the day.

C. Analyses of Reflectance Data

27. A fairly simple performance analysis technique was also used for the reflectance data. First, an appropriate natural background material (grass) was selected. Second, reasonable bounds were determined for the solar, near IR, and visual reflectances of grass. These bounds were determined by selecting the minimum and the maximum grass reflectances measured. The bounds established were as follows: solar reflectance, 19 to 34 percent; near IR reflectance, 27 to 37 percent; visual reflectance, 5 to 9 percent. The third step in the analysis process was to examine the reflectances of the materials to assign them into two categories, those materials judged to have camouflage potential and those judged to be camouflage deficient.

28. Materials with camouflage potential had to meet three requirements. Their average reflectance had to be within the established reflectance bounds of the background, their average reflectance had to be higher than that of the reference for the solar and/or near IR bands and lower for the visual band. These requirements lead to materials which have camouflage potential in both the thermal and the visual portions of the spectrum. Materials with high solar and near IR reflectances will usually absorb less of the sun's energy. Thus, they will stay cooler and be good for thermal camouflage. Materials with low visual reflectance will reflect less of the incident sunlight back to the observer and hence will be good for visual camouflage. Table 2 is the summary table for the aircraft shelter reflectance data. From the table one can see that only the Dutch low-emissivity paint was judged to have overall camouflage potential. Generally, the other materials were judged deficient because their solar and near IR reflectances were too low.

D. Analyses of Contact Temperature Data

29. The analysis method used for the contact temperatures was nearly the same as that used for the radiometric temperatures. However, there were several differences. The temperature of the background material was not grass temperature, but rather air temperature measured at each site. Air temperature is a reasonable approximation of natural background temperatures (e.g. trees and grass). For the contact temperature devices, 48 data points per day were collected and contact temperatures were collected on 55 days. For each data point (temperature sample), an average temperature of the material being examined was computed by adding the

Table 2

Spectral Reflectance Performance of Thermal Camouflage Materials

Material	Solar Performance	Near IR Performance	Visual Performance	Overall Performance
Dutch Low Emissivity Paint	+	-	+	+
Reference	-	-	-	-
Jute Camouflage Net	-	-	-	-
British Camouflage Net	-	-	+	-
British Low Emissivity Screen and Net	-	-	+	-

Note: + denotes camouflage potential
- denotes camouflage deficiency

temperature readings from 30 consecutive one-minute samples together and dividing by 30. Next, the absolute value of the temperature difference between the camouflage or reference material and its appropriate background material (air temperature) was calculated. Two measures of camouflage effectiveness were calculated based on these absolute value calculations. The first of these measures was a daily sum of the absolute values for each camouflage measure and each reference surface. This provides a measure of the total daily temperature differences between the camouflage or reference material and its background. The second of these measures was closely related to the first measure. The daily sum for each camouflage material was divided by the daily sum for its appropriate reference material and then this quotient was subtracted from one. This measure compares the camouflage and reference materials. If the result of this calculation is one, then the camouflage is perfect. However, if the result of this calculation is less than zero, then the camouflage material is less effective than the untreated reference surface.

30. Table 3 shows the contact temperature results. Only 4 of the 55 days are shown in the table. These days were carefully chosen from the entire data set to span the weather conditions experienced during the 1980 field trial. Data were selected from the sunniest, the cloudiest, the coldest, and the wettest days. From the table one can see that the best camouflage measure for both MOE#1 and MOE#2 was the wet jute camouflage net. However, the wet concrete and the British LE Screen also provided an improvement in thermal camouflage over the untreated reference surface.

Table 3

Contact Temperature Measures of Effectiveness (MOE) for Camouflage Materials for the Aircraft Shelter

Camouflage or Reference Material	Sunniest 18 Sept 80 MOE#1	Sunniest 18 Sept 80 MOE#2	Wettest 11 Oct 80 MOE#1	Wettest 11 Oct 80 MOE#2	Coldest 3 Nov 80 MOE#1	Coldest 3 Nov 80 MOE#2	Cloudiest 5 Nov 80 MOE#1	Cloudiest 5 Nov 80 MOE#2	Summary MOE#1	Summary MOE#2
South Wall										
LE Paint (Netherlands)	294	-0.2	48	-0.8	203	-0.3	46	-0.5	591	-0.3
Reference Wall	246	0	27	0	162	0	31	0	466	0
Wet Jute (U.S.) Camouflage Net	88	0.6	21	0.2	114	0.3	41	-0.4	264	0.4
Wet Concrete	116	0.5	37	-0.4	143	.1	36	-0.2	332	.3
Mat (Denmark)	385	-0.6	62	-1.3	283	-0.7	81	-1.7	811	-0.7
LE Screen (U.K.)	263	-0.1	39	-0.5	100	0.4	41	-0.4	443	0.1

Note: MOE#1 = $\Sigma|C-B|$ and MOE#2 = $1 - (\Sigma|C-B|)/(\Sigma|R-B|)$

where C = temperature of camouflage material
B = temperature of background (grass or trees)
R = temperature of untreated reference surfaces

IV. Concluding Remarks

31. This paper has presented an overview of Project TIREX as well as a sampling of some of the results for one of the structures on which camouflage materials were tested. The examples presented were limited to analyses of contrast reduction potential in both the visual and thermal IR bands. The results are indicative of the results of the complete data set. Few materials exist that offer an ability to completely reduce the contrast of a structure to its natural surround. While many materials offer significant thermal or visual contrast reductions, none were ideal for simultaneously providing the desired reduction in both bands. This indicates a continuing need to develop better materials and applications concepts.

32. While the results of ground-based data analyses may appear discouraging, when coupled with the initial imagery analyses the picture is not so bleak. Structures that are not extremely hot targets, such as aircraft shelters and hardened concrete structures, are detectable primarily by their individual shapes. The clever use of many of the existing materials to accomplish shape disruption can be very effective in reducing the detectability of those structures. The ground-based data provide a valuable quantitative data base to aid in the design of material application concepts. These concepts will allow use of these materials for modifying the appearance of structures to reduce their vulnerability to attack by thermal IR devices. Combinations of materials may provide significant reduction in the vulnerability of structures to both visual and thermal IR target acquisition aids. Insulating materials can have a significant impact on the detectability of hot targets such as the heated corrugated metal building investigated.

33. Project TIREX will provide the first available systematic guidance for thermal camouflage design. While a considerable amount of developmental work is still required, a very substantial foundation will be available to proceed rapidly to the full camouflage capability needed to address the modern threat.

AN APPROACH TO RISK AND UNCERTAINTY IN BENEFIT-COST ANALYSIS OF WATER RESOURCES PROJECTS

AMBROSE GOICOECHEA, Ph.D.
International Water Resources Institute
School of Engineering and
Applied Sciences
George Washington University
Washington, DC

MICHAEL R. KROUSE
Institute for Water Resources
U. S. Army Corps of Engineers
Fort Belvoir, Virginia

LLOYD G. ANTLE, Ph.D.
Institute for Water Resources
U. S. Army Corps of Engineers
Fort Belvoir, Virginia

1. INTRODUCTION

Traditionally, many benefit-cost studies in water resources planning have been conducted under the assumption that the "most likely" or the "expected" values of benefits and costs provide an appropriate economic basis for the evaluation and comparison of alternative project plans. It is not always acknowledged that estimates of those benefits and costs can vary over a range of values, and that study recommendations ought to reflect, somehow, the likelihood of estimates over their respective ranges. Even when the effort is made to consider parameter ranges, it is not clear as to how to proceed with the analysis—what analytical tools are to be used, and what tradeoffs to generate in order to assist in the decision-making process.

President Carter reinforced the need to develop systematic procedures for the analysis of risk and uncertainty in his directions to the Water resources Council (WRC) on June 12, 1978 to produce guidelines under which Federal Water Resources agencies conduct benefit-cost analysis. As a result, the WRC outlined a project evaluation procedure (May 24, 1979) to assist planners deal with aspects of risk and uncertainty. In that procedure it was suggested that "a range of reasonably likely outcomes can be described by using sensitivity analysis—the technique of varying assumptions as to alternative economic, demographic, environmental and other factors, and examining the effects of these varying assumptions on outcomes of benefits and costs." Furthermore, it was suggested to " ... characterize the range of outcomes with a set of subjective probability estimates. Where it is appropriate to do this, methods for manipulating the subjective probability estimates can then be used. However, if numerical estimates of subjective probability are given, it should be made clear in the project report that these are subjective and thus do not have the same objective basis that characterizes a long hydrologic record ...". The outline went on to suggest some general concepts to account for differences in risk and uncertainty among project purposes and costs, time periods, and stages of planning.

Responses to the above guidelines deemed the effort unsatisfactory, as noted in the "Federal Register of December 14, 1979 (pg. 72896). The principal thrust of the comments was that, "while it provided a clear statement of the problem and basic principles of risk and uncertainty in water resources planning, it failed to provide a usable procedure." The procedure as written was judged too general, permissive rather than directive, and lacking in instructive detail. Commenters noted "the omission of accepted mathematical and probabilistic methodologies and the failure to identify and list key factors to which benefit-cost evaluation is sensitive."

This heightened awareness of the role of risk and uncertainty in water resources planning motivated the work reported in this paper. Specifically, we have three objectives: (1) to raise some issues and concerns about the possible impacts of risk and uncertainty considerations on the preparation and documentation of feasibility studies, the development of alternative plans, the complexity of the decision-making process, and the time required for that process to evolve, (2) to identify sources of risk and uncertainty in water resources planning, (3) to investigate the computational requirements of a probabilistic approach to benefit-cost analysis, as well as its potential usefulness (or lack of it) in the evaluation of alternative projects.

2. SOURCES OF RISK AND UNCERTAINTY

A distinction is made in defining the terms "risk" and "uncertainty," (Water Resources Council, May 24, 1979, pg. 30196). Situations of risk are defined as those in which the random nature of potential outcomes can be described in reasonably well-known probability distributions. In situations of uncertainty, on the other hand, probability distributions are not known and their estimation would be a predominantly subjective exercise. However, the distinction between risky and uncertain situations is not usually a crisp one; rather, it is a matter of degree.

Risk is usually associated with our understanding of physical phenomena for which the probability of occurrence of an outcome can be estimated in a relatively objective manner (i.e., through the use of hydrologic records and similar data banks). When the assertion is made that an outcome will take place with a probability of P, the term risk is used to denote the probability (1-P) of other outcomes. Uncertainty, on the other hand is more closely associated with physical phenomena, socio-economic interactions, environmental response, pollutant effects on health, and other situations where our understanding is limited. These phoenomena and events are not inherently random, but they are relatively poorly understood. The term uncertainty is used, then, to reflect the subjective nature of the probabilities when these are assessed.

Incorporation of risk and uncertainty considerations in the decision-making process, then, calls for the identification of sources of R & U, reduction of R & U whenever possible, the generation of benefit-cost-risk tradeoffs, and public involvement. Risk can be reduced in some instances by collecting additional data and using the full length of available historical records (e.g., hydrologic records), increasing safety factors in design (which increase costs, in turn), selecting system performance measures with better-known characteristics, and by using more refined analytical techniques.

3. EXISTING AGENCY PROCEDURES FOR R & U

Water resources agencies, in general, allow for elements of risk and uncertainty in the decision-making process. This process often extends over several decades involving several feasibility studies, numerous requests for additional information, and the scrutiny of a chain of decision makers (agency review groups, political representation, and the general public).

In the case of the U.S. Army Corps of Engineers (which is of particular interest to the authors) the planning process starts with an assessment of needs, and these are reassessed many times before the project is constructed, and even many times after it becomes operational. Planners do not have much difficulty in documenting existing conditions and problems, but it becomes extremely difficult to characterize future conditions as well as the benefits and costs associated with those conditions. Projection of trends into the future become a complex and troublesome task. They are complex because they require an attempt to model and relate economic, social, demographic and political interactions over extended periods of time. Also, these interactions are frequently simplified to keep the model (e.g., computer program, or set of analytical tools) manageable, thus further increasing the uncertainty inherent in the model projections.

Agency review procedures are aimed at sorting out the conditions which influence the design of alternative plans for a project and that have a major impact on benefit-cost (B/C) ratios. When a project report clears preliminary reviews at Corps field divisions and at Office of the Chief of Engineers (OCE), the Board of Engineers for Rivers and Harbors (BERH) staff proceeds to review the project recommendations, project analysis, assessment of needs, and adherence to institutional policy. At this stage in the review, special emphasis is placed on the analysis of safety aspects of design criteria, and whether a project recommendation is supported by competent B/C analysis. Projects that exhibit a $B/C > 1.0$ will tend to be favored generally (there are other non-economic criteria to consider, as well, as will be discussed later) over those projects with a $B/C \leq 1.0$.

If project justification is close (e.g., B/C is only slightly greater than 1.0), a determined examination of sources of risk and uncertainty may be made to discern whether the recommendation is likely to survive long-run trends toward higher discount rates, and whether demographic and economic projections are consistent with trends forecast by the Office of Business Economics Research Service (OBERS). In addition, recommendations by the Chief of Engineers, the Secretary of the Army, and the Executive Office of the President may request that the project be subjected to restudy before advanced engineering funds are approved.

As stated early in the paper, the need to develop systematic procedures for the analysis of R & U has been perceived, however, and it may be appropriate now to inquire about the nature of the additional information to procure and document. Effective ways of displaying, "digesting" and integrating that information into the decision-making process become important also, particularly at the levels of (1) field division, (2) Board of Engineers of Rivers and Harbors, and (3) public participation.

4. PROBILITY DISTRIBUTION

The procedure presented in the next section requires that a probability distribution be computed for each benefit and each cost component of the economic analysis. Such a distribution function would be defined over the range of values of a particular component, and it would be characterized by an average or mean value, and a variance. Once the various probability distributions have been generated, the next step is to aggregate them statistically into some measure of economic performance (e.g., net benefits, B/C ratio, other).

For many problems there are available sufficiently long historical records to estimate probability distribution or relative frequencies of occurrence of the input parameters of interest. This is true particularly of input parameters that relate to the risk portion of the analysis - number of rainfall events in a given region within a year, intensity and duration of rainfall events, amount of runoff, etc. These distributions would be of an objective nature, based on available records, and fairly reliable. However, probability analysis with many economic or demographic variables such as population projections, price elasticity, demand for recreation activities, etc., often depends substantially on subjective estimates. Subjective probability distributions are thus inferior to objective probability distributions but invariably preferable to the option of omitting a measure of uncertainty altogether.

The suggested procedure begins with three estimates of possible values for each input component (e.g., benefits and costs): an *expected value,* an *optimistic value,* and a *pessimistic value.* The expected value is the value that is likely to occur, on the average, most often if the activity (e.g., flood event, extent of recreation use, time required to complete a task, etc.) could be repeated a number of times under similar circumstances. An optimistic value is the maximum benefit (or minimum cost) that could be realized if "everything went right" — schedules are met, costs of building materials are accurately assessed, demand for recreational uses materializes, etc. Finally, a pessimistic value is the minimum benefit (or maximum cost) that could be realized if "everything went wrong." Effectively, then, the optimistic and pessimistic values identify the range of possible values for a given input component. Additionally, it is suggested in this procedure that the range representsa "95 percent confidence interval," i.e., 95 percent of the time the actual value would fall within the specified range.

Once the three parameter estimates are prepared, it is possible to use them to estimate the "mean" and the "variance." There remains, then, the matter of specifying the "shape" of the distribution to be able to talk about the probability of that input parameter realizing a particular value within its range.

Gamma Distribution

A candidate is the gamma distribution of the form:

$$f(x) = \frac{x^{\alpha-1} e^{-x/\beta}}{\Gamma(\alpha)\beta^{\alpha}}, \qquad \text{for } 0 < x, \tag{1}$$

$$= 0, \text{ elsewhere}$$

where α and β are the parameters that determine the general shape of this function. The first two moments of this function are:

$$\text{mean,} \qquad E(\underset{\sim}{X}) = \mu = \alpha\beta \tag{2}$$

variance, $$E(\underset{\sim}{X}-\mu)^2 = \sigma^2 = \alpha\beta^2 \tag{3}$$

and a typical graph is shown below in Figure 1.

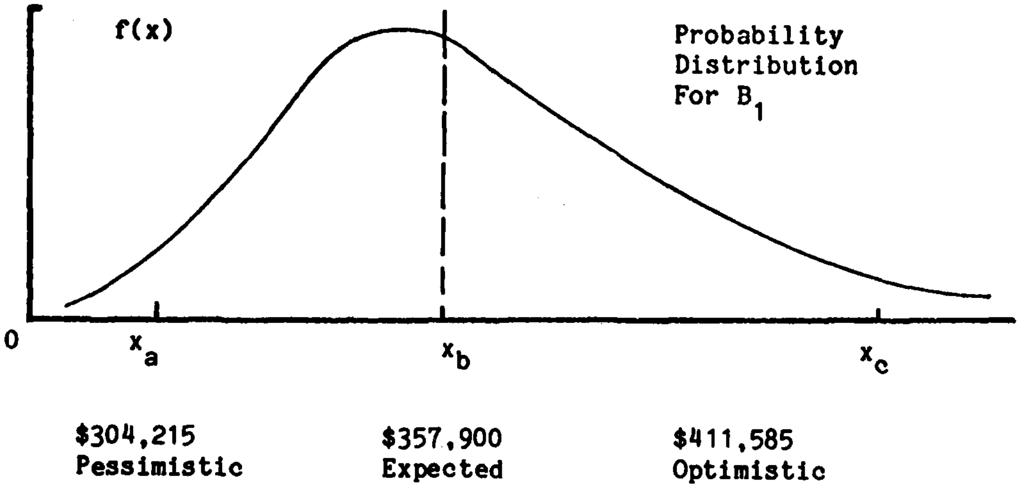


Figure 1. Gamma Distribution

One approach is to let $\mu = x_b$ and to approximate the variance using the relationship

$$\sigma^2 = \left[\frac{1}{6}(x_c - x_a)\right]^2 . \tag{4}$$

When used in this manner, the advantage of the gamma over the weibull distribution is that it requires only two parameter values, α and β, to specify it completely. In addition, and most important, the mathematical attributes of the gamma distributions make it possible to aggregate random variables analytically, as will be shown.

Normal Distribution

Another candidate is the normal distribution of form

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad \text{for } -\infty < x < \infty \tag{5}$$

where the parameters μ and σ^2 are the expected value and variance, respectively.

In contrast to the other proposed random variables, the normal has a range of values from $-\infty$ to $+\infty$ which renders it less suitable for benefit-cost analysis where input components generally have non-negative values, only. Obviously, the reason for our considering it as a candidate for analysis is the convenience of mathematical manipulation that it affords. In particular, sums of normal random variables, are normally distributed. For this distribution it is also suggested that

$$\mu = x_b \qquad (6)$$

and

$$\sigma^2 = \frac{1}{6}\,(x_c - x_a)^2 \qquad (7)$$

There are other candidate distribution functions potentially useful in benefit-cost analysis, and our intent was to identify but a few.

5. ANALYTICAL PROCEDURE

We now proceed to investigate the analytic (as opposed to the Monte Carlo, computer simulation approach) feasibility of generating a probability distribution of the B/C ratio itself. First, we will present the main theoretical foundation of the procedure.

LEMMA 1. Let $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_n$ be statistically independent normal random variables with parameters (μ_i, σ_i^2) for i = 1, 2, ..., n. Then, the new random variable $\underset{\sim}{Y} = \underset{\sim}{X}_1 + \underset{\sim}{X}_2 + \ldots + \underset{\sim}{X}_n$ is normally distributed with parameters $\mu = \mu_1 + \mu_2 + \mu \ldots + {}_n$, and $\sigma^2 = \sigma_1^2 + \sigma_2^2 + \ldots + \sigma_n^2$.

LEMMA 2. Let $\underset{\sim}{X}_1, \underset{\sim}{X}_2, \ldots, \underset{\sim}{X}_n$ be statistically independent gamma random variables with parameters (α_i, β_i) for i = 1, 2, ..., n. Then the new random variable $\underset{\sim}{Y} = \underset{\sim}{X}_1 + \underset{\sim}{X}_2 + \ldots + \underset{\sim}{X}_n$ has the moment generating function of form

$$M(t) = \left[(1 - \beta_1 t)\alpha_1 (1 - \beta_2 t)\alpha_2 \cdots (1 - \beta_n t)^{\alpha_n} \right]^{-1} \qquad (8)$$

for all real t such that $\beta_i t < 1$, where $\beta_i = \max \left\{ \beta_1, \beta_2, \ldots, \beta_n \right\}$.

The above results for lemmas 1 and 2 can be verified easily using the definition and unique properties of the moment generating function.

LEMMA 3. Let $\underset{\sim}{X}_B$ and $\underset{\sim}{X}_C$ be two statistically independent normal random variables with parameters (μ_i, σ_i^2) for i = B, C. Then, the new random variable $\underset{\sim}{Y} = \underset{\sim}{X}_B/\underset{\sim}{X}_C$ has the distribution

$$g(y) = Y \left[4\Pi^2 \sigma_B^2 \sigma_C^2\right]^{-1/2} \int_{-\infty}^{\infty} \exp \left[-\frac{1}{2}\left(\frac{zy - \mu_B}{\sigma_B}\right)^2 + \left(\frac{z - \mu_C}{\sigma_C}\right)^2\right] dz \qquad (9)$$

for $- \infty < Y < \infty$. The proof is given in Goicoechea et al. (1982).

LEMMA 4. Let $\underset{\sim}{X}_B$ and $\underset{\sim}{X}_C$ be two statistically independent gamma random variables with parameters (α_i, β_i) for i = B,C. Then the new random variable $\underset{\sim}{Y} = \underset{\sim}{X}_B/\underset{\sim}{X}_C$ has the distribution

$$g(y) = \frac{(\alpha_B + \alpha_C) y^{(\alpha_B + \alpha_C)}}{\Gamma(\alpha_B)\ \Gamma(\alpha_C) \beta_B^{\alpha_B} \beta_C^{\alpha_C}} \left[\frac{y}{\beta_B} + \frac{1}{\beta_C}\right]^{-(\alpha_B + \alpha_C)} \qquad (10)$$

for $0 < y_1$, and zero elsewhere.

A general procedure that makes use of the lemmas above is now outlined to generate a distribution function of the new random variable of interest, $\underset{\sim}{B}/\underset{\sim}{C}$. Let

$$\underset{\sim}{Y} = \frac{\underset{\sim}{B}}{\underset{\sim}{C}} = \frac{\underset{\sim}{B}_1 + \underset{\sim}{B}_2 + \cdots + \underset{\sim}{B}_n}{\underset{\sim}{C}_1 + \underset{\sim}{C}_2 + \cdots + \underset{\sim}{C}_m} \qquad (11)$$

where $\underset{\sim}{B}_i$ is the random variable representing the i-th benefit component, and $\underset{\sim}{C}_i$ is the random variable representing the i-th cost component. In this approach it is suggested that the analysis be performed with gamma variables only, or normal variables only for analytical convenience.

Gamma variables only,

1. for the i-th benefit component, set the estimate of the expected value, X_b, equal to the parameter μ, e.g., $\mu_i = (X_b)_i$,
2. compute the variance ${}_i$ using the relationship

$$\sigma_i^2 = \frac{1}{6} (X_{ci} - X_{ai})^2$$

3. set $\mu_B = \mu_1 + \mu_2 + \cdots + \mu_n$, where n is the number of benefit components.
4. set $\sigma_B = \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2$.
5. perform steps 1 through 4 above for each of the m cost components to compute μ_c and σ_c^2,
6. substitute the values of μ_B, μ_C, σ_B^2, and σ_C^2 in the expression for $g(y) = g(B/C)$, equation (10)
7. evaluate g(y) numerically at various values of y and plot the results.

Normal variables only,

1. for the i-th benefit component, set the estimate of the expected value, X_b, equal to the parameter μ, e.g., $\mu_i = (X_b)_i$,
2. compute the variance σ_i using the relationship

$$\sigma_i^2 = \frac{1}{6}(X_{ci} - X_{ai})^2$$

3. compute the parameters α_i and β_i such that

$$\beta_i = \sigma_i/\mu_i, \text{ and } \alpha_i = \mu_i/\beta_i$$

4. set $\alpha_B = \alpha_1 + \alpha_2 + \cdots + \alpha_n$, where n is the number of benefit components,
5. estimate β_B such that

$$(1 - \beta_B t)^{\alpha_B} = (1 - \beta_1 t)^{\alpha_1} (1 - \beta_2 t)^{\alpha_1} \cdots (1 - \beta_n t)^{\alpha n}$$

6. perform steps 1 through 5 above for each of the m cost components to compute α_C and β_C
7. substitute the values of α_B, α_C, β_B and β_C in the expression for $g(y) = g(B/C)$, equation (9),
8. evaluate g(y) numerically and plot the results.

The reader will note that the analysis assumes that the random variables above are statistically independent. In actuality the physical activities that they represent (e.g., recreation, employment, operation and maintenance, etc.) will be related to each other.

6. AN ILLUSTRATIVE EXAMPLE: THE HALSTEAD FLOOD PROTECTION PROJECT

The city of Halstead is located in the Little Arkansas River Basin, which covers an area of about 1,400 square miles. Periods of heavy rainfall cause widespread flooding in and around the city, and structural measures had been considered for reducing flood damage. In a recent study (U.S. Corps of Engineers, 1976), a flood control levee and a channel improvement project, as well as added recreation were proposed to remedy the

TABLE 1. Estimate of range of values for the Halstead (Kansas) flood protection project.

Costs ($)

	Optimistic Value	Expected Value	Pessimistic Value
1. Interests and Amortization	197,250	263,000	328,750
2. Operation and Maintenance	14,175	18,900	23,625
Total:	211,425	281,900	352,375

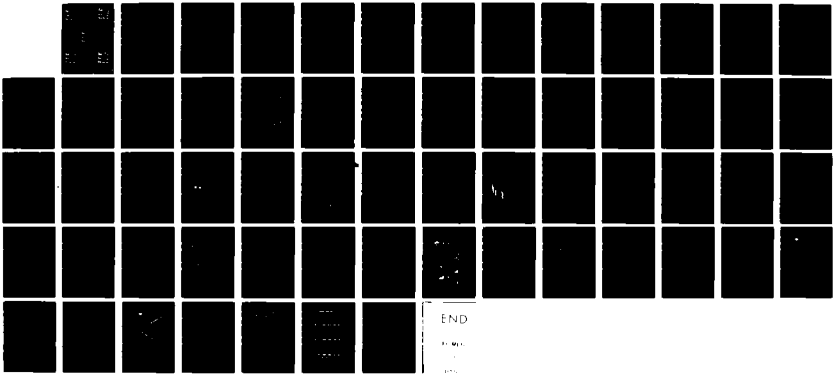
Benefits ($)

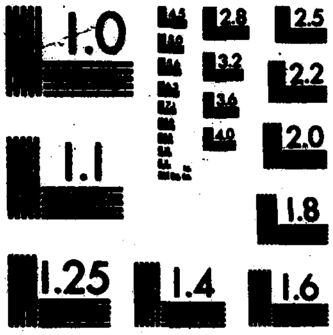
	Optimistic Value	Expected Value	Pessimistic Value
1. Inundation Reduction	411,585	357,900	304,215
2. Location	51,810	47,100	42,390
3. Affluence	26,840	24,440	21,960
4. Floodproofing Costs Prevented	16,000	8,000	0.0
5. Employment	39,800	18,400	0.0
6. Recreation	7,500	2,600	0.0
Total:	533,535	458,440	368,565

	1.04	low
B/C =	1.63	medium (expected)
	2.62	high

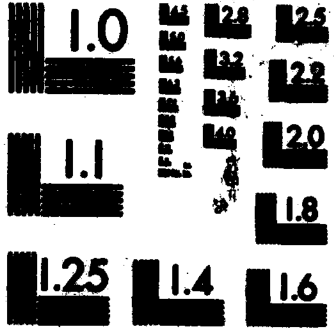
AD-A120 811 PROCEEDINGS OF THE 1982 ARMY SCIENCE CONFERENCE HELD AT THE UNITED STATES..(U) DEPUTY CHIEF OF STAFF FOR RESEARCH DEVELOPMENT AND ACQUISITIO.. 18 JUN 82 6/6

UNCLASSIFIED F/G 5/2 NL

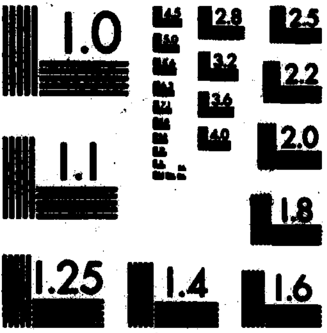



1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

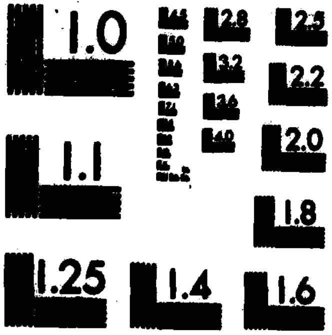
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

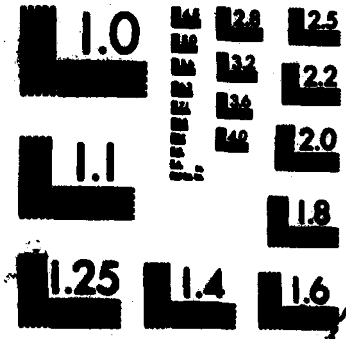
MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

problem. The study report indicated annual costs of $281,900, and annual benefits of $458,440, as shown in the expected value column of Table 1, yielding a benefit cost ratio of 1.6.

To investigate the potential usefulness of the proposed approach, as well as its limitations, a series of interviews were conducted with a water resources planner who was thoroughly familiar with the study. Considering each cost and benefit component, one at a time, optimistic and pessimistic values were estimated as shown in Table 1. As it turned out, the elicitation of estimates was not particularly painful because current practice is such that documented costs (expected values) already carry a 25% increase "contingency factor". The planner, also an experienced engineer, thought it appropriate and felt comfortable about doing away with that 25% increase to estimate the optimistic value, and add a second 25% increase to produce the pessimistic value. This was done for the two cost components.

For the benefits, the planner felt that each component had to be considered separately, on its own merits, and that whereas the risk analysis portion of the study allowed him a measure of confidence, the uncertainty associated with projections (e.g., population, employment, recreation) tended to widen the range of his estimates. A factor of 15% was offered for inundation reduction, followed by 10% for both location and (benefits resulting from increases in the amount and quality of furnishings, appliances, and inside content of households in general). Again, each time the assigned percentage is subtracted from the expected value to generate the pessimistic value, and added to generate the optimistic value. However, when it came to recreation benefits the planner felt that uncertainty considerations played a predominant role and opted for a pessimistic value of zero benefits and an optimistic value of $7500—a 100% and over variation on both sides of the expected value of $2,600. Similar considerations were applied to the estimation of employment and floodproofing benefits. Needless to say, the above estimates reflect the professional experience and attitude towards risk and uncertainty of this particular planner.

The general procedure previously outlined was then applied to "aggregate" the estimates elicited above. This was done first assuming normal distributions over those estimates then with gamma distributions, yielding the results shown in Figure 2. Several observations are in order. Steps 1 through 6, for the normal assumption, were readily performed but evaluation of $g(y)=g(B/C)$ required numerical integration for subsequent display. The numerical integration was a non-trivial exercise and required the use of a computer program to deal effectively with a wide range of z values (see the integral in equation 9), and with the product of small numbers. With the gamma assumption, on the other hand, evaluation of $g(y)=g(B/C)$ required no integration. Instead, the numerical evaluation involved the manipulation

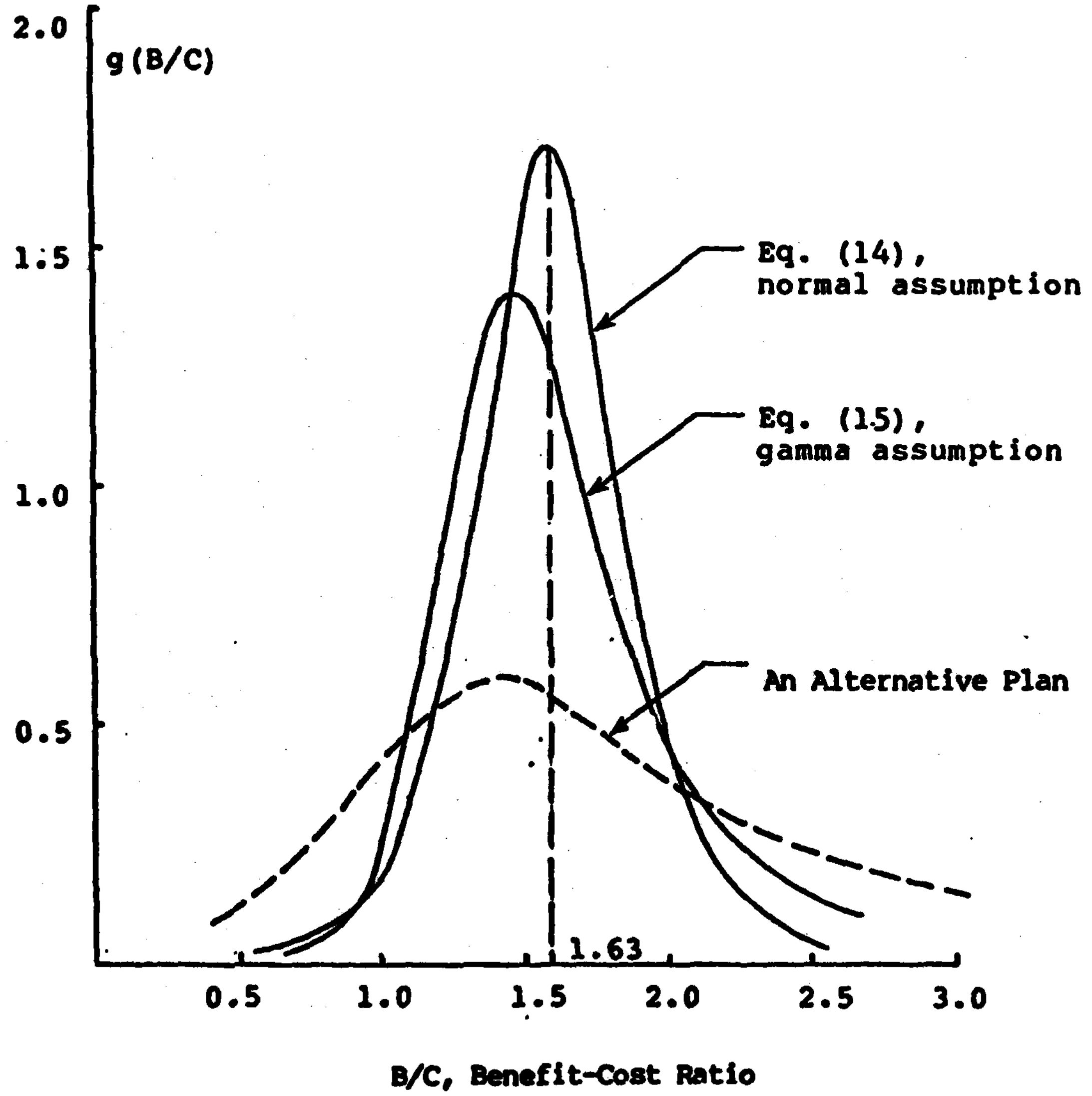


FIGURE 2. Computed Distribution of the B/C Ratio.

of very large numbers corresponding to values of the gamma function $\Gamma(.)$. Again, a computer program was required for that purpose.

As Figure 2 shows, a B/C ratio of 1.63 corresponds to the expected value of the random variable Y = B/C. This result was anticipated, but there was no way of anticipating the actual shape of its distribution by just looking at the estimates presented on Table 1. It is this distribution of the B/C ratio that can now be used to state the probability that the B/C ratio will be, say, less than or equal to a certain number. This probability value, of course, is equal to the area under the distribution function for values of B/C less than and equal to that certain number of interest, as made available by the cumulative distribution function F(B/C). For the gamma assumption, for example,

$$\text{PROB } [B/C \leq 1.63] = 0.48$$
$$\text{PROB } [B/C \leq 2.00] = 0.80$$

A point to make, then, is that B/C = 1.63 is not a fixed value and that 48% of the time the actual value will be less or equal to 1.63. Conversely, 52% of the time the actual value will be equal to or greater than 1.63.

7. CONCLUSIONS

Several questions arise when the implementation of this procedure (or any other) is considered. How can the information generated so far be of use to a decision-maker? Should a distribution of the B/C ratio be obtained for each alternative plan identified in a feasibility study? Does the additional information generated justify the additional expense and documentation needed? Will it shorten or lengthen the planning process? And, ultimately, will it contribute to making better decisions? These are difficult questions to answer, and experience with the proposed general procedure and similar efforts would be needed before definitive assessments can be made. However, we can observe that:

1. an estimated range of values for each benefit and cost component may be a reasonable item of information to elicit from planners; ideally, these ranges are arrived at after the uncertainty analysis is applied to the evaluation of the param[illegible] of the flood's p.d.f., as well;

2. given two alternative plans with the same expected B/C ratio, but different variances, the plan with the smaller variance should be preferred over the other plan; in Figure 2, for instance, the "alternative plan" (with the dashed line distribution) is clearly an inferior plan (on Pareto efficiency grounds);

3. given two alternative plans with different expected B/C ratio values; but equal B/C ratio variances, the plan with the larger expected B/C ratio value should be preferred over the other plan;

4. for a finite number of alternative plans, each one with a potentially different combination of B/C expected value and variance, it is not immediately apparent how to proceed with a ranking and selection of a preferred alternative. The analyst may have to resort to multiobjective analysis to explore the tradeoffs (see Goicoechea et. al., 1982, for a description of multiobjective ranking techniques),

5. benefit-cost ratio, or net benefits is not the only criteria for project evaluation and selection; in fact current procedures call for the attainment of national economic development (NED), environmental quality (EQ), social well-being, and regional economic development objectives; the needs exist, it would appear, to consider benefit-cost analysis within a multiobjective framework; in that framework, the B/C expected value and variance represent useful information in a list of criteria to use to evaluate and rank alternative plans.

6. the matter of arriving at estimates of the ranges appropriate for the various components of benefit and cost, can be difficult at times and can vary from one planner (e.g., decision-maker) to another; by necessity, these estimates are of a subjective nature and reflect whether a planner is risk-averse, risk-prone, or risk-neutral; however, once the estimates are arrived at, the analytical procedure developed can be applied to aggregate the effects of parameter variation and produce a probability distribution of the B/C ratio.

REFERENCES

- Goicoechea, A., D. R. Hansen, and L. Duckstein, Multi-Objective Decision Analysis with Engineering and Business Applications, to be published by John Wiley and sons in 1982, 640 pp.

- U.S. Corps of Engineers, "Survey Report for Halstead Local Flood Protection Project," Feasibility Report, September 1976.

- U.S. Water Resources Council, Procedures for Evaluation of National Economic Development (NED) Benefits and Costs in Water Resources Planning (Level C), Federal Register, Vol. 44, no. 102, pg. 30194-30258, Washington, DC, May 24, 1979.

A COMPUTER-BASED INTERACTIVE MODEL FOR INDUSTRIAL LAND USE FORECASTING

AMBROSE GOICOECHEA, Ph.D.
International Water Resources Institute
School of Engineering and
Applied Sciences
George Washington University
Washington, DC

MICHAEL R. KROUSE
Institute for Water Resources
U. S. Army Corps of Engineers
Fort Belvoir, Virginia

1. INTRODUCTION

An industrial engineering activity that is growing in relevance and receiving due attention in the literature is that of identifying land areas suitable for future industrial use. As cities expand and multiply, the various activities that reflect the social-economic makeup of a community (e.g., industrial, commercial, residential, agricultural, etc.) compete with each other for use of the same fixed resource—land. It then becomes necessary and meaningful to consider the science and art (e.g., economic and behavioral aspects) of land use forecasting.

Land use forecasting has long been a planning activity of interest to the various Federal and State agencies, particularly those with mandates for the development of land and water projects. Certainly this is the case at the U.S. Army Corps of Engineers, where land use forecasting has long been applied to the evaluation of economic benefits resulting from engineering measures and associated land uses. Over the last 50 years a number of research efforts have been funded by the Corps relating to the development of analytical land use methodologies and, in some cases, the design of computer-based forecasting models.

The purpose of this paper, then, is to review briefly the progress made in the analytical and behavioral development of land use forecasting models, to point to the modeling functions of special relevance to industrial land uses, and describe a new interactive computer model being developed at the Institute for Water Resources (IWR) of the U.S. Army Corps of Engineers.

2. BACKGROUND

A substantial number of efforts to develop land use forecasting models have been undertaken over the last 30 years. This section compiles a list of over 50 models created during that time period that cover a wide range of forecasting activities, and that represent the extent of the modeling effort in the private and public sectors.

The beginnings of land use forecasting in the United States are to be found in the schools of city planning created at Harvard University in 1929, and at the Massachusetts Institute of Technology (MIT) in 1931. These two schools spearheaded the tremendous development that resulted thereafter. Kilbridge et al.[1] presents a classification of 20 urban planning models by land use (e.g., industrial, commercial, residential, agricultural, etc.), function (e.g., projection, allocation, and derivation), theory (e.g., behavioral, gravity, trend, and growth index), and method (e.g., regression, input-output, markov process, linear programming, and simulation) that span the time period 1959-1967. Table 1 extends and updates that classification by identifying 28 other models that are considered most significant and that cover the time period (1962-1979). As Table 1 reveals, these models offer a wide range of forecasting capabilities, use diverse analytical and behavioral approaches, and have been applied to a good number of cities in the U.S. Also, as the reader can observe, industrial land use forecasting is well integrated and represented in many of these models (see refs. 2 through 29 and 38).

The remainder of this section reviews some of the previous work that led to the development of the Alternative Land Use Forecasting (ALUF) model of the Institute for Water Resources, U.S. Army Corps of Engineers.

Land Use Forecasting at the Corps of Engineers. Essential to the task of project development and evaluation is the determination of "with project" and "without project" future economic conditions. The calculation of these economic benefits has provided, in fact, the motivation for much of the effort on land use forecasting at the Corps. But substantial and continuous as this effort has been over the last decade, the need still exists to develop a computer package that offers a satisfactory balance of sound methodological framework, data base and computer time requirements. Some of the methodologies proposed in the past, although analytically correct and based on sound methodological frameworks, were inadequate for subsequent implementation for several reasons. At times the theoretical development was valid and well researched, but the computer model was incomplete. More frequent was the case, however, where the proposed

Table 1. A Classification of Land-Use Forecasting Models

#	MODEL NAME	PROJECT NAME/ CASE STUDY	FORECASTING CAPABILITIES	ANALYTICAL TOOLS EMPLOYED	REFERENCE	DATE
1	BASS-Bay Area Simulation Study	San Francisco Bay Area	• Employment (5 year inter.) • Population • Industrial Land Use (21) • Residential Housing Location • Gov. Employment	• Regression • Judgmental Weighting	Center for real estate and urban economics, University of California, Berkely Also, H.J. Brown (1972) Ref. (2)	1962 - 1968
2	PLUM - Projective Land Use Model	San Francisco Bay Area	• Household Location • Population-Serving Employment • Industry Location • Land Use • Regional Employment • Regional Population	• Regression Functions • Subjective Probability	W. Goldner (1968) (3)	1968
3	Puget Sound Regional Transportation Study	Puget Sound Regional Plan-Commission Seattle, Wash.	• CBD Employment • Industrial Location • Population • Retail • Pop. by county • Employment by industry	• 16 Sector Input-Output Model • Judgmental Weighting	C.H. Grave (1964) (4)	1964 - 1970
4	SEWRPC	Southeastern Wisconsin Regional Planning Commission	• Regional Employment • Land Use Residential Industrial (L.P.) Agriculture	• Linear Programming • Judgmental • Input-Output	S. Wisconsin Regional Planning Commission Tech. Ref. 3 (5)	1966
5	TALUS	Detroit Regional Transportation and Land Use Study	• Employment by District • Households by District • Land Use • Employment • Population	• Regression • Gravity Access	Rubin, J.J. (1968) (6)	1968
6	HARVARD	Southwest Sector of the Boston Region, Exper. Study (16 week)	• Industrial Location • Residential. • Recreation, Open Space • Transportation	• Computer Simulation • Universal Transverse Mercator (UTM) Grid • Grid Network • Overlay	Steinitz and Rogers (1970) Proposal for Year Four (1974) Alzo, I.L. McHarg (1969) (7),(8)	1970 - 1974
7	MANGROVE	Rookery Bay Land Use Studies, Collier County Florida	Environmental Planning Strategies: • Canal Concept • Resource Buffer • Filling of Welands • Land Preservation	• Ecological System Modeling • Judgmental	Center for Urban Studies, Miami University, FL. A.R. Veri et.al (1973) (9)	1973
8	AIRPORT	Airport Environs: Land Use Control	• Coordinated Report of Land Use Planning Controls, Noise Reduction		Office of Metropolitan Planning and Development, Environmental Planning Division Washington, DC (10)	1970

9	FMS	Floodplain Management System Model; Rillito River, Tucson, Arizona	o Land Use Allocation o "With and without" Analysis o Incremental Analysis o Economic Analysis o Population Distribution	o Linear Programming o Regression	IWR Paper 74-82 WEISZ and Day (1974) (11)	1974
10	UCM	St. Louis Region	o Determine Demand for industrial location in flood plain.	o Factor o Discriminant Analysis o Statistical Analysis	IWR Paper 74-P8 Corbean and Meyer (1974) (12)	1974
11	1 DYLAM-Dynamic Land use allocation model	Cleveland Study,	o Population o Employment o Land use o Changes in Infrastructure	o Grid Network o Proximity Factors o Graphic Display	Seader and Grava (1971) (13)	1971
12	OAK RIDGE LAB.	6500-sq mile Region in Eastern Tennessee	o A cell-based land-use model o Data base of pop dist, labor available, size & freq of industries o Cumulative distribution of existing industries o Employment by zone.	o Factor analysis o Attractiveness scores o Statistical analysis o Delphi techniques	P.A. Leaslie (14)	1976
13	EMPIRIC	Boston Area (Plus a dozen other cities)	o Population by Zone o Employment by Zone o Land use	o Linear Difference equations o Statistical analysis o Non-behavioral	Peat et. al. (1971) (15)	1971
14	Harvard Model	The interaction between urbanization, Land Quality and quality	o update vacant land values		Bloom and Brown (16)	1979
	Landscape Architecture Research Office, Harvard University (See Attached computer readout)	- Land value model				
15		- Housing Model	o Residential Land Use by single and multiple-family structure.		Wilkins et. al. (17)	1979
16		- Public institutions model			Vidal and Brown (18)	1979
17		- Transportation model	o Travel Demand on Transportation Facilities		Tyler and Commings (19)	1979
18		- Industrial Model	o Industrial use Siting Based on: Slope Depth to bedrock Zoning o Development costs used as economic criteria		Coltry et. al. (20)	1979
19		- Public expenditures model	o Land use Change based upon public local expenditures		Kirlin et. al. (21)	1979

20	Harvard (cont.)	- Water Quantity and Quality	o Water demand o Coliform Count o Carbonaceous BOD o Dissolved Oxygen o Salinity		Rogers and Berwick (22)	1978
21		- Vegetation-wildlife Model	o Produce land classification systems o Resource Evaluation o Description Methods		Smith (23)	1979
22		- Commercial model	o A developer's perspective in estimating the location and size of commercial centers		Wilkins and Brown (24)	1979
23		- Solid Waste Management	o Landfilling and export-out-technologies on a town-by-town basis	o Regression o Land use exclusion criteria	Rogers and McClellan (25)	1979
24		- Legal/Implementation model	o Land use allocation in accordance with state, federal, and local land use controls		Giezentanner, Steinitz (26)	1978
		- Historical Resources Model	o Evaluates the specific, unique, often Qualitative values of areas and buildings		Steinitz and Haglund (27)	1978
26		- Recreation Model	o Identifies sites suitable for recreational development and ranks them.		Steinitz and Douglas (28)	1978
27		- Soils Model	o Identification of soil erosion zones. o transport of sediment to accumulation areas o Estimation of costs associated with mitigating procedures			
28		- Land use Descriptors	o Cover characteristics o Construction o Cost		Way, D.S. (29)	1978

computer model required vast amounts of input data, the exogeneous parameters themselves were difficult to estimate (e.g., spatial population distributions) or the amount of time required to apply the model would have been unreasonably large (in the order of months).

An alternative course of action is delineated here. Essentially, some of the computer subroutines in program RIA are combined with an economic data bank file and a search procedure to allocate land uses. Optimal land allocations are not sought; instead, "near optimal," feasible land allocations are desired.

3. ALTERNATIVE LAND USE FORECASTING (ALUF) PROGRAM

The development of a grid cell data file requires that each variable map be individually encoded and geographically registered to a common base and stored, along with data variables in the data bank, on a computer storage device.

The IWR package consists of two computer programs which are used in connection with a grid cell spatial data base as shown in Figure 1. The main program, Alternative Land Use Forecasting (ALUF) does the actual allocation of future land uses to specific grid cells. The Existing Land Use Analysis Program (ELUA) is provided to help identify significant land use location factors for the allocation process based on the relationship between land use locations and other data available in the grid cell data bank.

The final program output is a new data variable written into the data base file for each grid cell, indicating projected future land use. The programs are written in FORTRAN IV for the CDC 6600/7600 series computers.

The ALUF program incorporates the HEC RIA Attractiveness modeling program and RIA Distance Determination package. These were adapted for use in this process so that land use locator scores can be developed according to user specified criteria, as well as location criteria derived from the statistical findings.

The kinds of data variables commonly used as a basis for allocating future land use include:

A. Access (Distance)
    1. Transportation
    2. Central Business Districts or Regional Centers
    3. Dependent Activities
B. Proximity to Compatible Land Uses
C. Physical Land Attributes (Developability)
    1. Slope
    2. Drainage
    3. Type of Cover
    4. Soils
D. Infrastructure
    1. Sewers and Water
    2. Gas and Power
    3. Mass Transit
E. Zoning
F. Ownership
G. Land Prices

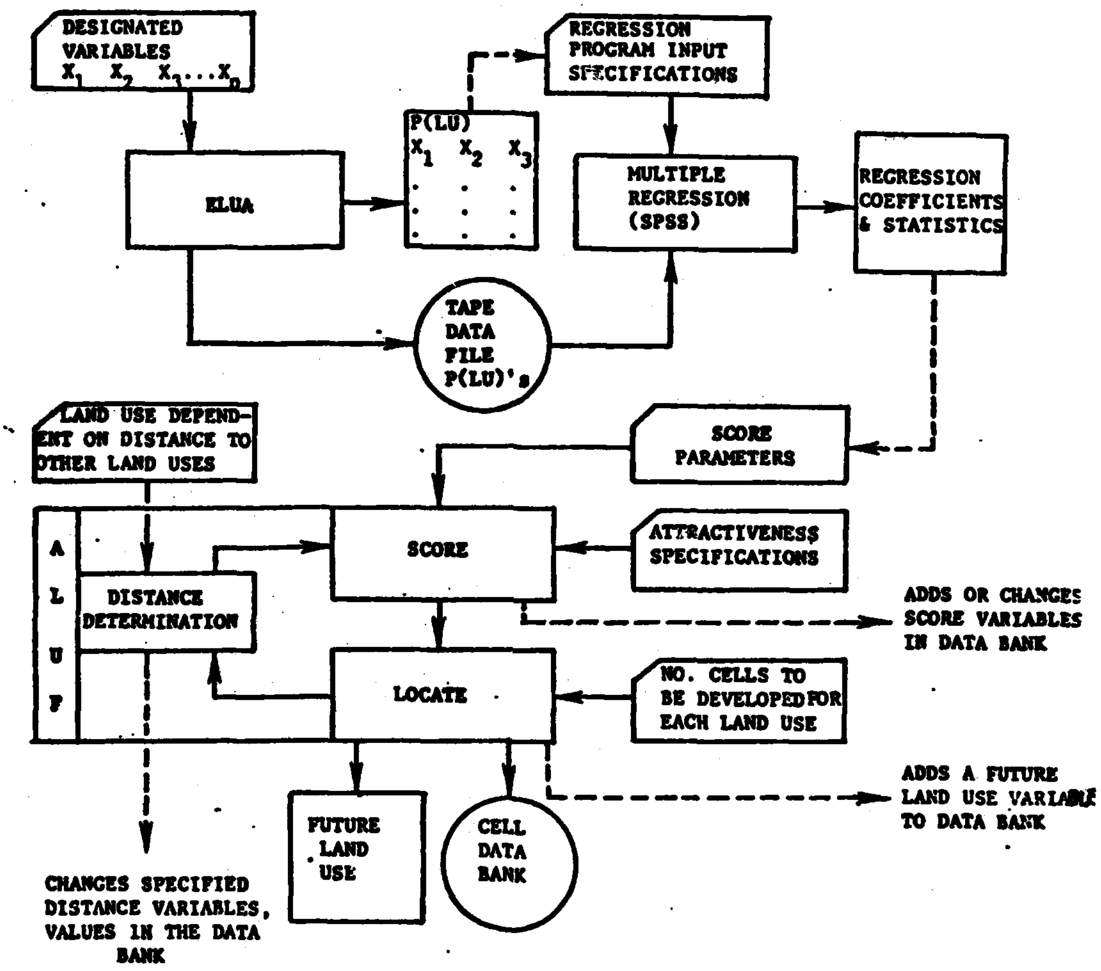

DESIGNATED VARIABLES X_1 X_2 $X_3 \ldots X_n$
ELUA
P(LU) X_1 X_2 X_3
REGRESSION PROGRAM INPUT SPECIFICATIONS
MULTIPLE REGRESSION (SPSS)
REGRESSION COEFFICIENTS & STATISTICS
TAPE DATA FILE P(LU)'s
LAND USE DEPENDENT ON DISTANCE TO OTHER LAND USES
SCORE PARAMETERS
A L U F
SCORE
ATTRACTIVENESS SPECIFICATIONS
DISTANCE DETERMINATION
ADDS OR CHANGES SCORE VARIABLES IN DATA BANK
LOCATE
NO. CELLS TO BE DEVELOPED FOR EACH LAND USE
ADDS A FUTURE LAND USE VARIABLE TO DATA BANK
FUTURE LAND USE
CELL DATA BANK
CHANGES SPECIFIED DISTANCE VARIABLES, VALUES IN THE DATA BANK

Figure 1. Current Structure of Computer Program ALUF

4. CALCULATION OF THE ATTRACTIVENESS INDICES

To demonstrate the computation of the raw attractiveness indices as performed in the computer program ALUF, consider the land use grid shown in Figure 2. Purposely, the grid is small and contains only 27 cells, so as to render the exercise workable (typically grid representatives of regions of interest may require 5,000-50,000 cells).

Listed in Figure 2 is the legend used to represent the various land uses, e.g., (1) natural vegetation, (2) developed open space, (3) low density residential, etc. In this manner, we can see that grid cell (i,j) = (1,4) is currently allocated to low density residential. A railroad track traverses the grid network, as shown.

As program ALUF is structured currently, a matrix arrangement is available to the analyst to identify the variables (topographic) of interest, as shown in Table 2. The analyst-user then is required to: (1) designate topographic variables, (2) assign relative weights to the variables, and (3) specify a shading intensity for each value of each designated variable. A matrix must be filled in for each land use (e.g., activity) being considered. For illustrative purposes, Table 2 alone is shown with the matrix values for industrial use.

We continue our illustrative computation of the attractiveness indices for industrial use with the specification of two variables only: (1) distance to Seaboard Railroad (variable #23), and slope (variable #8). Information on these two variables must be built into the data bank file prior to running the program. For our example, this information would appear as shown in Figures 3 and 4. With reference to location (i,j) = (1,4), we notice that the slope value of 2 corresponds to a "2 to 6 percent slope" (Table 5, variable 8, Appendix), and the distance to the railroad tracks is three cell units. The actual computation of the raw Attractiveness Index proceeds as follows:

Distance to R.R. = 3 : (1) x (2) = 2
Slope = 2 : (8) x (1) = $\underline{8}$
10

Shading Intensity
Relative Weight
Attractiveness

In a similar manner, indices (also called scores) for the remaining cells are computed in Table 3 and again shown in Figure 5.

There remains the matter of using the attractiveness scores to allocate a land use to each grid cell. Currently, the program assigns land uses according to the priority identified by the analyst in the Data Deck;

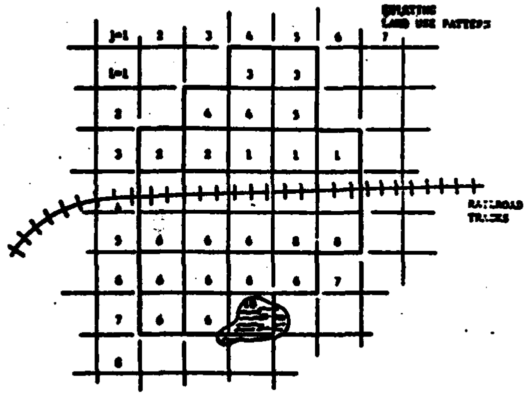

EXISTING LAND USE PATTERN
j=1
i=1
RAILROAD TRACKS

Figure 2. Existing Land Use Pattern

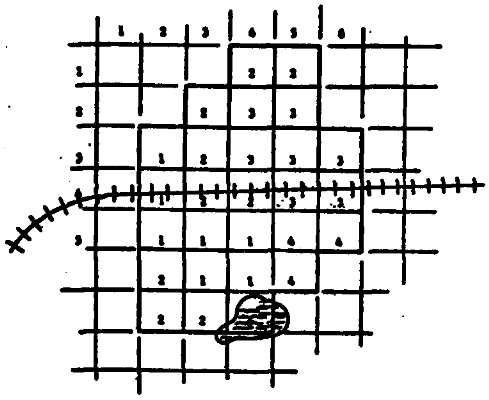


Figure 3. Slope Associated with Grid Cells (Variable 8)

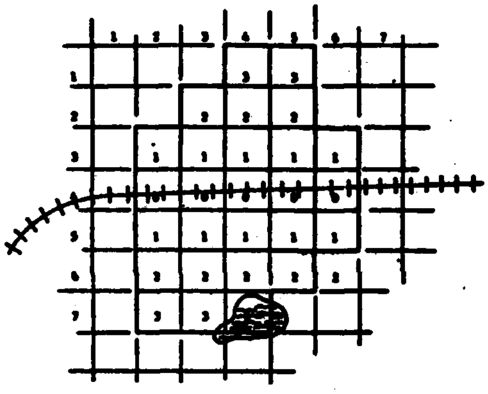


Figure 4. Distance in Cell Units to Railroad Tracks (Variable 23).

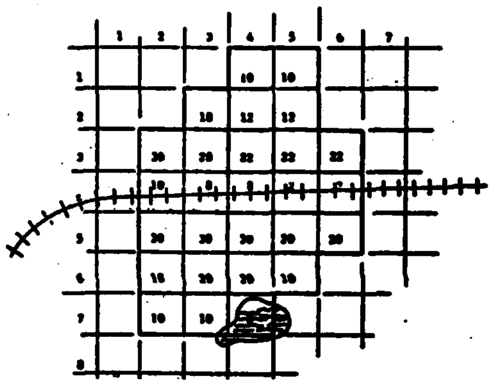


Figure 5. Raw Attractiveness Scores for Industrial Use

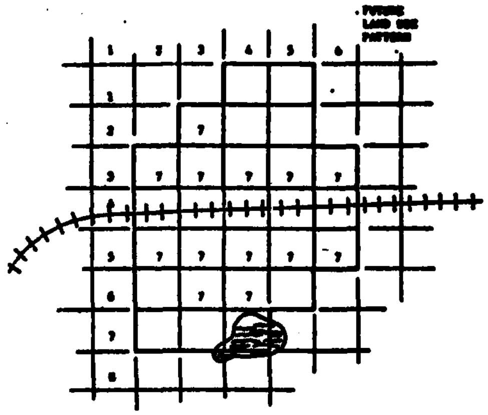

FUTURE LAND USE PATTERN

Figure 6. Future Industrial Land Use

LEGEND:

1. Natural Vegetation
2 Developed open space (parks, golf, ...)
3 Residential, Low Density
4 Residential, Medium
5 Residential, High
6 Agricultural
7 Industrial
8 Commercial
9 Pasture
10 Water Bodies

Table 2. Attractiveness Matrix for Industrial Use

	Topographic Variable	Shading Intensity 0	1	2	3	4	5	Importance Weight
(23)	Distance to R.R.	0	10	5	1	0		2
(8)	Slope	-1	10	8	2	0		1
	Existing Land Use	0	10	-1	-1			
(17)	Distance to E.R.		0					
	Distance to E.I.		10	10	9			

NOTE: Shading intensity values range from -1 to 10. A value of 10 is assigned if variable is of most significance to land use being considered. A value of -1 is assigned if variable is to be excluded completely from further consideration.

Table 3. Attractiveness Scores for Industrial Use

CELL NO.			SCORES		
(i,j)	DIST R.R.	SLOPE	DIST R.R.	SLOPE	TOTAL
1,4	3	2	(1)(2)=2	(8)(1)=8	10
1,5	3	2	(1)(2)=2	(8)(1)=8	10
2,3	2	2	(5)(2)=10	(8)(1)=8	18
2,4	2	3	(5)(2)=10	(2)(1)=2	12
2,5	2	3	"	"	12
3,2	1	1	(10)(2)=20	(10)(1)=10	30
3,3	1	2	(10)(2)=20	(8)(1)=8	28
3,4	1	3	(10)(2)=20	(2)(1)=2	22
3,5	2	3	"	"	"
3,6	1	3	"	"	"
4,2	0	1	0	(10)(1)=10	10
4,3	0	2	0	(8)(1)=8	8
4,4	0	2	0	(8)(1)=8	8
4,5	0	3	0	(2)(1)=2	2
4,6	0	3	0	(2)(1)=2	2
5,2	1	1	(10)(2)=20	(10)(1)=10	30
5,3	1	1	"	"	30
5,4	1	1	"	"	30
5,5	1	4	(10)(2)=20	(0)(1)=0	20
5,6	1	4	"	"	20
6,2	2	2	(5)(2)=10	(8)(1)=8	18
6,3	2	1	(5)(2)=10	(10)(1)=10	20
6,4	2	1	(5)(2)=10	(10)(1)=10	20
6,5	2	4	(5)(2)=10	(0)(1)=0	10
7,2	3	2	(1)(2)=2	(8)(1)=8	10
7,3	3	2	"	"	10
7,4	3	0	(1)(2)=2	(-1)∴REJECT	-

that is, if the desired priority is industrial, followed by high density residential, low density residential, commercial, etc., then the analyst physically places data cards for industrial at the top of the "Data Deck," followed by data cards for high density residential, and so on. In that manner, given a request for 13 cells, say, for industrial, the program assigns a land use (Legend Code 7) to the 13 cells that exhibit the highest industrial attractiveness score. A similar allocation rationale is then used for high density residential, and so on down the priority list. For our example then, the cells allocated to industrial use are shown in Figure 6. Note that for cell(6,5) there corresponds a slope value of 4 (i.e., 10 to 15 percent grading) and that Table 3 shows a shading intensity of zero; the slope variable, then, contributes a value of zero to the attractiveness score, e.g. (0)(1.0) = 0.0. Cell(7,4), on the other hand, has a slope value 0.0 (i.e., water body) and since an intensity value of -1 has been assigned to it, the cell is excluded from industrial use.

5. AN ILLUSTRATIVE COMPUTER APPLICATION: TRAIL CREEK PILOT STUDY

Now that the computation of the raw attractiveness scores has been illustrated in a step-by-step manner, the application of the procedure to a real-world situation is demonstrated using computer program ALUF. The region of interest is the Trail Creek study area shown in Figure 7, and it exhibits variety and complexity of roads, railroad track, river lengths, urban center nearby, etc. Current land use of this area is as shown in Figure 7, with adopted dimensions for each rectangular cell of 200 and 333.3 feet.

The interactive computer mode of the program was then used to fill in the attractiveness matrices. This time it is noted that the exercise was extended beyond the industrial land use stated requirement to include residential and commercial. The number of cells required for each use was 900, 800 and 200, respectively.

Finally, shown in Figure 8 is the computer printout of the computed future land use pattern. Only the left half of the pattern is used, as the other half would be of a similar nature. The actual computer printout does yield the two halves, however. Let us now compare existing and future land use of a particular cell, say cell(35,55). It is observed that Figure 7 identifies the current use as being agricultural (i.e. code number 6), and now the future use is projected to be industrial (i.e. code number 7), as given in Figure 8.

6. SUMMARY AND CONCLUSIONS

This paper discusses the architecture and use of a new land use forecasting model labeled ALUF, Alternative Land Use Forecasting. The model makes use of information on current land uses, topographic characteristics, and preferences elicited from the planners to forecast future land uses.

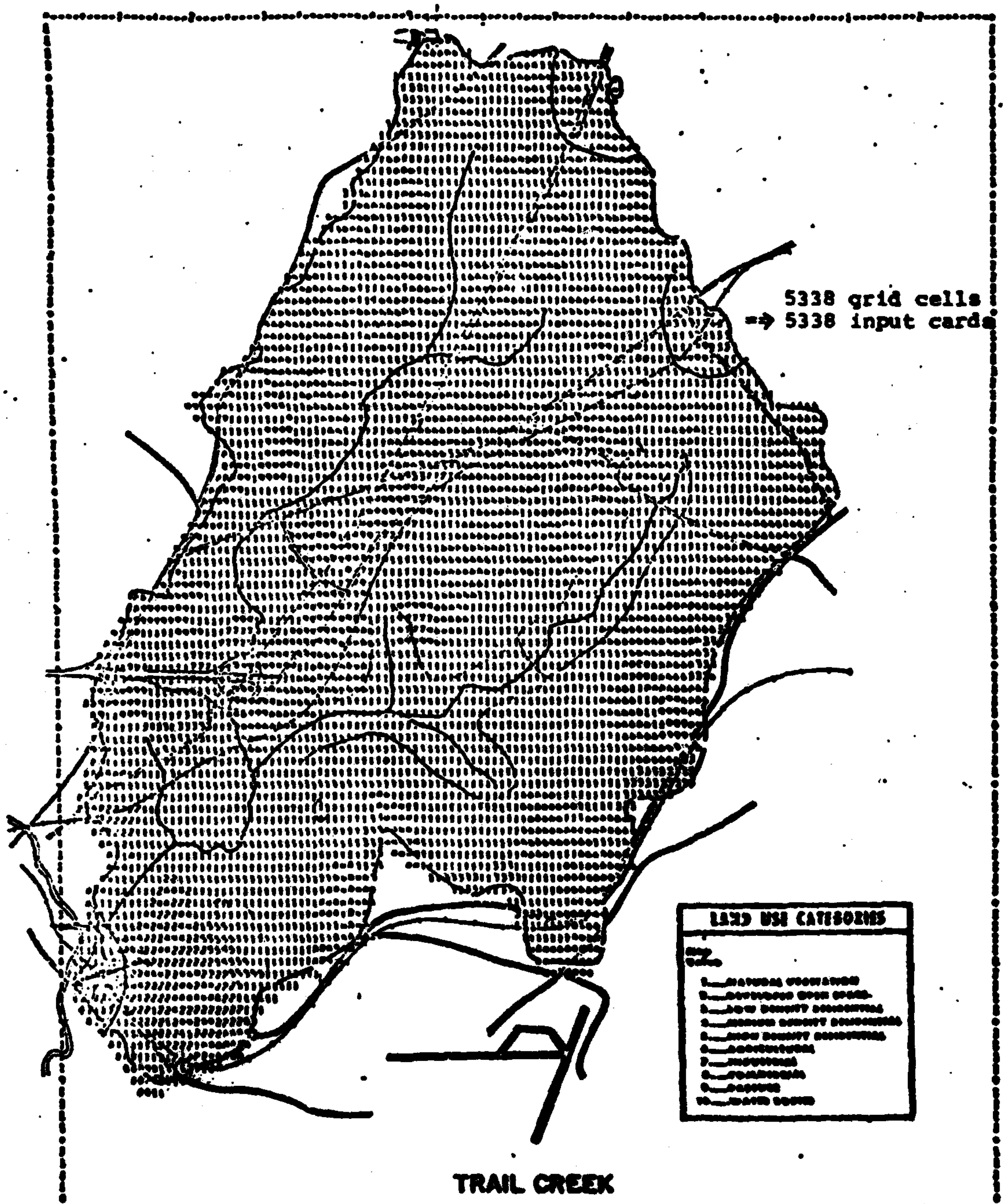
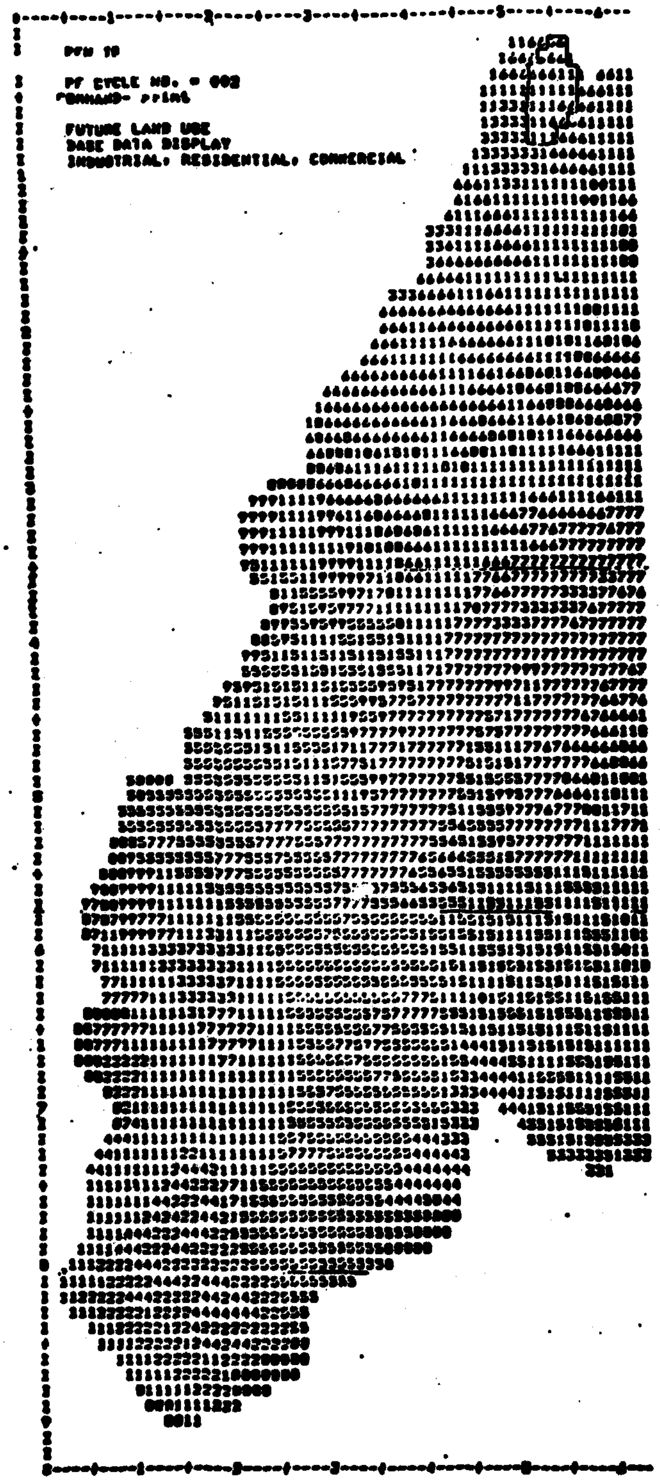

5338 grid cells
=> 5338 input cards
LAND USE CATEGORIES
TRAIL CREEK

Figure 7. EXISTING LAND USE PATTERN



```
PFN 10

PF CYCLE NO. = 002
COMMAND- print

FUTURE LAND USE
BASE DATA DISPLAY
INDUSTRIAL, RESIDENTIAL, COMMERCIAL
```

Figure 8. Computer Printout of Future Land Uses

In the process, it calculates the economic benefits to be derived from a proposed engineering measure or zoning policy. The model is currently operational and it is available to Corps personnel and general city planners involved in project development and evaluation. Also, it is hoped that industrial engineering practitioners will find it useful in their dialogue with city planners as new industrial enterprises in growing communities are discussed.

REFERENCES

Kilbridge, M., R. O'Block, and P. Teplitz, "A Conceptual Framework for Urban Planning Models," Management Science, Vol. 15, No. 6, 1969.

Brown, H. C., J. R. Ginn, F. J. James, J. F. Kain, and M. R. Straszheim, Empirical Models of Urban Land Use: Suggestions on Research Objectives and Organization, Exploratory Report 6, National Bureau of Economic Research, New York, 1972.

Goldner, W., Projective Land Use Model (PLUM), Bay Area Transportation Study Commission, BATSC Technical Report 219, San Francisco, California, September 1968.

Hydrologic Engineering Center, Resource Information and Analysis Using Grid Cell Data Banks, Computer Program 4D1-X6-L7590; User Manual, U.S. Army Corps of Engineers, Davis, California, September 1978.

Hydrologic Engineering Center, An Investigation of Concepts and Methods for Broadened Scope Flood Plain Information Studies, Phase I , Oconee Basin Pilot Study, U.S. Army Corps of Engineers, Davis, California, September 1975.

Jolissaint, C. H., Issues Related to Interfacing Water Resource Planning and Land Use Planning: Development and Application of Quantitative Procedures, INTASA, Inc., Menlo Park, California, May 1977.

Note — Space limitations preclude a complete listing of references. Please write to authors requesting such listing.

MICROCLIMATE COOLING FOR COMBAT VEHICLE CREWMEN (U)

RALPH F. GOLDMAN, Ph.D. and STAFF
U.S. ARMY RESEARCH INSTITUTE OF ENVIRONMENTAL MEDICINE
NATICK, MA 01760

I. INTRODUCTION

At air temperatures above 28°C (82°F), human thermal comfort is best achieved without clothing; however, it may be essential if the human is exposed to contact with surfaces above 45°C (113°F), sunlight, blowing sand, drying wind, missiles and/or NBC agents. Increasing protection is associated with increasing thickness of barrier layers on the skin. Unfortunately, any barriers between the skin and the ambient environment reduce the ability to eliminate heat from the body.

a) BODY HEAT PRODUCTION

Heat production (M) at rest is about 1 Met (50 kcal of per square meter of body surface area per hour); an average man's heat production at rest is 90 kcal/hr or 105 watts. Normal work can double heat production and hard work can triple it. The sustainable "voluntary hard work" level is ~5 Met (425 kcal/hr or 500 watts), while 6 or 7 Met will exhaust the average man if sustained for a few hours.

b) BODY HEAT LOSS MECHANISMS

About 12% of the resting heat production is eliminated from the lungs by respiration. Another 12% is eliminated by evaporation of the body water diffusing through the skin; up to 0.6 kcal of heat is removed for each gram of water evaporated. Evaporation can, however, be blocked by ambient vapor pressures greater than the vapor pressure of water at skin temperature or by clothing which, even if permeable to water vapor, reduces the potential evaporative cooling by imposing insulation between the skin and the environment. The remaining 76% of resting metabolic heat production is eliminated from the body by convection and radiation in a comfortable environment. Elimination of the increased heat production during work is facilitated by the convective air motion generated by body "pumping", but in a warm environment most of the increased heat production is lost by production of sweat and its subsequent evaporation; even in a cold environment, about 42% of working heat production may be lost by evaporation.

c) STILL AIR INSULATION; THE "CLO" UNIT

Even without clothing, there is a barrier layer of still air next to the skin. This still air film acts as insulation against heat exchange between the skin and the ambient environment; without body or air motion this external air layer (Ia) provides 0.8 clo of insulation. One clo unit of clothing insulation is defined as allowing 5.55 kcal/m^2/hr of heat exchange by radiation and convection (H:RC) for each °C of temperature difference between the skin (at average skin temperature Ts) and ambient air temperature (Ta). Since the average man has 1.8 m^2 of surface area, his H:RC can be estimated as: H:RC = (10/clo) (Ts - Ta) [Eq. 1]; i.e. an 0.8 clo still air layer limits the heat exchange by radiation and convection for a nude man to about 12.5 kcal/hr (10/0.8) for each °C of difference between skin and air temperature. Thus, producing 90 kcal/hr, a resting man will lose 11 kcal/hr (12%) by respiration, 11 kcal (12%) by evaporation of the water diffusing through his skin and will have a requirement to evaporate sweat (Ereq) to eliminate the remaining 68 kcal/hr if the Ta is less than 5.5°C (i.e. 68/12.5 kcal/hr/°C) below Ts. The required sweat evaporative cooling (Ereq) can be estimated as: Ereq = M + (H:RC) [Eq. 2] where M is the heat produced during rest or work and H:RC is estimated by Eq. 1. Since a comfortable Ts is about 33°C (91.4°F), an increasing percentage of the body surface area will be required to be sweating with a Ta above 27.5°C (i.e. 33° - 5.5°).

The external air layer is reduced by air motion, approaching a minimal value of ~0.2 clo at air speeds above 4.5 m/s (10 mph). With this minimum air insulation (0.2 clo), 68 kcal/hr can be eliminated by a nude man at an air temperature only 1.4°C below skin temperature (i.e. 68/(10/0.2) = 1.36°) without sweating.

d) CLOTHING INSULATION

Studies of clothing materials have concluded that clothing insulation is a linear function of thickness; differences in fiber or weave, unless these affect thickness, have only minor effects on insulation. A typical value for clothing insulation is 1.57 clo per centimeter of thickness (4 clo per inch). Figure 1 displays the actual thickness of the intrinsic insulation layers found around three body segments (torso, arm and leg) with ordinary clothing; the contributions of the trapped air layers to the total thickness are far greater than the thickness contributed by the fabric layers. Even with foam materials, used in some protective ensembles, the trapped air between layers is the dominant factor in insulation. Since insulation is a function of thickness and this, in turn, is a function of the number of layers, each added layer of clothing will exert a characteristic increase in total insulation. Thus, most two layer clothing ensembles exhibit quite similar insulation characteristics regardless of differences in fiber, fabric type or layer thickness.

e) EVAPORATION THROUGH CLOTHING

Evaporative heat transfer through clothing also is limited by its thickness. The moisture permeability index (im) is a dimensionless unit with a lower limit value of 0 for an impermeable layer and an upper value of 1 if all the moisture that the environment can take up can pass through the fabric. Values of im approaching 1 are only found with high wind and no clothing, since moisture vapor transfer is limited by the characteristic value for diffusion of moisture through still air. An im value for typical clothing in still air is ~0.5. Water repellent treatments, very tight weaves and chemical protective impregnations reduce im significantly.

The evaporative heat transferred from the skin, through the clothing and external air layers, to the environment *is not simply a function of the permeability* index (im) but a function of the permeability index-insulation ratio (im/clo). The maximum evaporative heat exchange with the environment can be estimated, as in Eq. 1 for the H:RC of a man, as: $HE\ max = 10 \times im/clo \times 2.2\ (Ps - \O a\ Pa)$ [Eq. 3] where Ps is the vapor pressure of sweat (water) at skin temperature, Ts; Øa is the fractional relative humidity and Pa is the saturated vapor pressure at air temperature. Thus, the maximum evaporative transfer is a linear, inverse function of insulation even if not further degraded by specific chemical agent protective treatments which reduce permeability or by water repellent treatments.

f) THE "PHYSIOLOGIC PROBLEM" OF PROTECTIVE CLOTHING

The percent sweat wetted surface area (%SWA) is the ratio of the required evaporative cooling (Ereq) estimated by Eq. 2, to the maximum evaporative cooling ("Emax") estimated by Eq. 3; i.e. %SWA = Ereq/Emax [Eq 4] While a little sweating is not uncomfortable, as the body surface area wet with sweat approaches 20%, discomfort begins to be noted. Discomfort is marked with between 20 and 40% of the body surface sweat wetted and performance decrements can appear; they increase as %SWA approaches 60%. Sweat begins to be wasted, dripping rather than evaporating at 70%. Physiological strain becomes marked between 60 and 80% SWA; increases above that level limit tolerance even for fit, heat acclimatized men. Obviously, any conventional chemical protective clothing will pose severe tolerance limits since their im/clo ratios are rarely above 0.2. The basic problem is that skin temperature (Ts) must be maintained at least 1°C below deep body temperature (Tre) for the body to transfer enough heat from the body core to the skin.

Normally, under conditions of unlimited evaporation, skin temperature is about 3.3°C + (0.006° x M) below Tre. Thus at rest, when Tre is 37°C, the corresponding Ts is about 33°C. The 4°C difference between Tre and Ts allows each liter of blood flowing from the deep body to the skin to transfer 4 kcal of heat to the skin. Since Tre increases and Ts decreases with increasing M, it usually becomes easier to eliminate body heat with increasing work since the difference between Tre and Ts increases by about 1°C per 100 watts of increase in M. Thus, at a sustainable voluntary hard work level (M = 500 watts) each liter of blood flowing from core to skin can transfer 9 kcal to the skin, 225% more than at rest.

Unfortunately, any clothing interferes with heat loss from the skin and skin temperature rises, predictably, with increasing clothing. Core temperature (Tre) also rises when clothing is worn, as a function of the insulation induced rise in Ts and the resulting limited ability to transfer heat from core to skin. There is an even greater interference with heat loss from the skin when sweat evaporation is required (Ereq) but is limited either by high ambient vapor pressures (Øa Pa), low wind or low clothing permeability index (im/clo) (cf. Eq. 3). As Ereq approaches Emax, skin temperature increases dramatically and deep body temperature begins to increase exponentially. Deep body temperatures above 38.2°C are considered undesirable for an industrial work force; at a deep body temperature of 39.2°C, associated with a skin temperature of 36° or 37°C (i.e. $\bar{T}s$ converging toward Tre) there is a 25% risk of heat exhaustion collapse in fit troops. At an elevated $\bar{T}s$, and Tre of 39.5° there is a 50% risk of heat exhaustion collapse and as Tre approaches 40°, with elevated skin temperatures, almost all individuals are highly susceptible. Tre levels above 42°C are associated with heat stroke, a life threatening emergency.

g) GENERAL CONCLUSIONS AS TO NATURE OF THE PROBLEM

In essence, mission performance will be seriously degraded by CW protective clothing worn during heavy work in moderately cool environments, or at low work levels in warm conditions. Little reduction in heat stress is likely with any two layer protective ensemble, or any single layer vapor barrier system for protection against CW agents, unless some form of auxiliary cooling is provided.

Figure 2 is a 1963 chart of "Predicted Time to 50% Unit Heat Casualties" when troops wear a CW protective ensemble in either open (MOPP III) or closed (MOPP IV) state. This is expressed as a function of the environmental Wet Bulb Globe Temperature (WBGT) index. If hard work is involved, tolerance time to 50% heat casualties is between 1 and 2 hours, whether in MOPP III or MOPP IV, and almost without regard to ambient WBGT above 70°F. For moderate work, little problem would be anticipated with WBGT in the 70°F range for closed suit, or below 80°F for open suit. For light work the WBGT would have to reach 90°F for MOPP IV and about 97°F for MOPP III to incur 50% unit heat casualties in 5 to 6 hours.

II. A FIELD STUDY DEMONSTRATION OF THE PROBLEM

a) DESIGN OF THE STUDY

Several XM-1 tanks were available at Yuma, AZ to study the heat stress/CW protection problem under desert conditions. Two Marine tank crews volunteered as subjects; they were superbly fit, well-trained, heat-acclimatized and motivated. A six day test was carried out; the first two days were for training, and resolving problems; the last four days comprised the data generating portion of the trial. Day 3 involved wearing the fully closed CW clothing system over the CVC uniform (MOPP IV); the vehicle hatches were left open, but the engine and ventilators were shut off. MOPP IV was also worn on the final three days of the study; the engine and ventilators were shut off and all hatches closed on these three days.

b) ENVIRONMENTAL CONDITIONS

During the four days of actual test the temperature averaged 35°C (95° F) with 26% relative humidity; winds were from 4 to 13 knots, cloud cover between 13 and 30%. There was little build-up of tank interior temperature above ambient temperature, even during Days 4, 5 and 6 when the hatches were closed and ventilation shut off. On Days 1 to 3, when the hatches were open, there was only a small increase in interior humidity over exterior relative humidity, by about 10%; however, when the hatches were closed (Days 4, 5 and 6), the interior relative humidity rose dramatically.

One uses WBGT index to describe the heat stress. On Day 3 the inside WBGT was actually lower than that outside the vehicle because of the reduction of the solar load component of WBGT. However, on Days 4 through 6 there was a substantial, progressive increase in interior WBGT throughout the exposure inside the vehicle, reflecting the build up in humidity as the men's sweat accumulated.

c) ACTIVITY (HEAT PRODUCTION) LEVELS

Prior to the field study, it was predicted that if activity were limited to moderate work (~200 kcal/hr) only heat exhaustion would be incurred with a few hours in the closed hatch conditions. Accordingly, only one 3 to 5 minute fire mission was performed every thirty minutes. The Driver was essentially at rest; his estimated heat production was about 100 watts (i.e. under 90 kcal per hour). The Loader was doing less work than either the Commander or the Gunner; the latter, who had the most sustained work during the 3 - 5 minute fire mission each half hour, is estimated to have had a heat production of, at most, 225 watts (200 kcal/hr).

d) METHODS

Each afternoon the volunteers were weighed nude. Thereafter, all fluid intake was measured using pre-weighed canteens. $\bar{T}s$ and Tre and heart rate sensors were attached, the crew men were dressed in the uniform for the day and reweighed.

The tank was parked in the sun close to a building where all data collection equipment was located. The vehicle had been wired for temperature measurements, and a network of cables connected from the vehicle to the measuring and monitoring equipment. This allowed on-line data collection and processing of Tre, $\bar{T}s$, interior and exterior air temperatures (DB), wet bulb temperatures (WB) and 2 measures of heat stress, the standard WBGT (FSN #6665-00-159-2218) and a new "BOTSball" WGT (FSN #6665-01-103-8547). The data were continuously recorded and graphed on-line. Heart rate was measured at appropriate intervals using a standard EKG. The crewmen entered the vehicle fully dressed; once inside it would have been impossible for them to dress in the CW protective clothing.

e) PHYSIOLOGICAL RESULTS

The rectal temperature (Tre) mean weighted skin temperature (MWST), the air (DB) and wet bulb (WB) temperatures and the WBGT and BOTSball temperatures are presented for Days 3 to 6 in Figure 3. All men wore the MOPP IV configuration.

On Day 3, Tre stayed low, even though all ventilators were shut off, because the hatches were open. One can see ripples associated with the 5 minute fire missions, especially in the skin temperatures of the Gunner and Commander. The men had no difficulty completing the scheduled exposure.

On Day 4, with hatches closed a very different pattern emerges; although there is a relatively constant difference between the interior and exterior air temperature (DB), the interior wet bulb (WB) rises dramatically. Since the non-ventilated wet bulb makes up 70% of the WBGT and BOTs indices, both of these show steeply rising heat stress. The effects of this stress are immediately notable in the steeply rising skin temperatures, and more slowly responding but nevertheless increasing deep body temperatures. Within 30 minutes, we detected errors in the

Commander's directions for the fire mission and within the first hour he noted he was "making dumb mistakes". Water intake was strongly encouraged; up to 3 canteens an hour were ingested. Despite these attempts to maintain hydration and a high degree of motivation and leadership, after 80 minutes the Gunner slumped back in his seat, tore off his gas mask and indicated that he could not continue. The Commander and Gunner had voiced complaints for some time, felt chilled, were a little dizzy, but had not reached the criteria for removal; H.R. $\geq$ 180 b/m; Tre $\geq$ 39.5° or Ts $\geq$ Tre. They had continued despite increasing discomfort. Although not at the criterion for termination, this voluntary discontinuance by the Gunner was not capricious; his final heart rate was 178 beats per min.

Essentially similar exposure conditions prevailed on Day 5, but the men wore a vest supplied with cooled water; this removed heat at a rate of about 100 watts from each man. Although the interior environmental humidity build-up did occur, there was little or no rise in Tre; $\bar{T}$s were extremely low. The men completed the full exposure without difficulty, without error and without discomfort.

Day 6 was a repeat of the Day 4 exposure. The ambient conditions were milder (WBGT was 35°C on Day 4 versus 33.4°C on Day 6) so that it took longer, but again it was the Gunner who, at 124 minutes of exposure, was unable to continue. This voluntary intolerance occurred as his skin temperature converged to his deep body temperature (cf Fig 3). As on Day 4, there were fire command errors, and subjective discomfort and complaints beginning early and increasing throughout the exposure, but the men did their very best to complete the full exposure.

The heart rate is perhaps the best expression of the combined effects of work, environment and clothing on the crewmen. On Day 3, with open hatches and ventilators off, heart rates were relatively steady but the average for the 4 man crew was above 100 beats per min. In contrast, on Days 4 and 6 heart rate rose linearly, reaching a peak average of 150 beats per min after 80 minutes on Day 4 and about 135 beats per min at 124 minutes on Day 6. When auxiliary cooling was provided, the average heart rate of the group was less than when the hatches were open throughout the exposure on Day 3.

On Day 3, with hatches open but with the men in full MOPP IV configuration, the sweat evaporation was not substantially different than on Day 1 when only the CVC uniform was worn, but it was achieved at a much greater expense in sweat production. With the hatches closed on Days 4, 5 and 6, the evaporation was stringently limited; the ratio of the amount of sweat able to be evaporated per unit of production (E/P) clearly showed the relative inefficiency of sweat elimination of body heat for Days 4 and 6; these were in the 20 ± 20% range, in contrast to the 30 to 50% values with auxiliary cooling on Day 5 or with open hatches of Day 3, and the 60 to 80% of Day 1 with just the CVC uniform. Sweat rate on Day 4 averaged 2 kg (4 1/2 lb) an hour; sweat rate for the Loader and Tank Commander exceeded 3 kg (6 1/2 lb) during the 80 minute exposure. The demands for water to replace these sweat losses can be contrasted with the average 0.63 kg/hr (1.4 lbs/hr) of sweat produced when auxiliary cooling was available. There is a reduction in drinking water requirement of between 1 and 1.5 liters an hour with auxiliary cooling.

f) PERFORMANCE RESULTS

With the closed hatch conditions on Day 4, the men knew they were in troublehalfway through. At the end some felt they could only continue for another 10 or 15 minutes, except the Driver who was having little problem because of his low work rate. Estimated ability to perform was decremented by 25% for the Loader, and by more than 50 to 60% for the Gunner and Tank Commander. With auxiliary cooling on Day 5, the men had no problem completing, and felt they could continue for 3 to 4 hrs; there was little or no decrement in ability to perform the missions.

g) SUMMARY AND CONCLUSIONS FROM THE FIELD TRIAL

Significant heat stress was demonstrated, at a level to produce early performance decrements and, eventually, subjective inability to continue; this was fully supported by physiological data as being a valid endpoint for performance capability. This occurred under relatively modest ambient environmental heat conditions, and was most identifiable when the hatches were closed and the ventilators and blowers shut off. Auxiliary cooling using a water-cooled vest was clearly demonstrated as capable of alleviating the heat stress.

In conclusion, we have identified a clear mismatch between the ability of a crewman dressed in CW protective clothing and the simple demand that he perform an extremely light fire mission when ambient conditions (expressed as the WBGT) are in the 32 to 35°C (90 to 95°F) range. In this study, these occurred inside the XM-1 only when the hatches were closed and the blowers shut off.

III. EVALUATION OF POSSIBLE SOLUTIONS

When one identifies such a mismatch between the man's capabilities and the demands of his mission, there are generally only three classes of solution: 1) modify the man; 2) modify the clothing or equipment; and/or 3) modify the mission. Everything possible to improve the tolerance of crewmen had been done in this test; the men were fully heat acclimated, had good training and practice wearing chemical protection, excellent physical fitness, superb motivation and leadership, and drank as much water as possible. This leaves only the latter two classes of solution to deal with the problem. First, it should be possible to re-design the tank ventilation system to avoid the interior humidity (and potential temperature) build-up when the hatches are closed. The only other simple solution is to revise tactics so as to minimize any closed hatch, ventilators off operation, limiting duration of such conditions to not more then 30 minutes. As with a revision of the tank ventilation system, this solution approach will not solve the heat stress problem globally, but it will reduce the range of environments in which it will be experienced. A suitable, and in some ways more functional, alternative is to provide some form of auxiliary cooling directly to the crewmen. A properly designed system will eliminate heat stress, conserve large amounts of drinking water and allow unimpaired performance across any climatic range, even in the Arctic if provision is made for heating the heat transfer medium.

IV. LABORATORY STUDIES ON AUXILIARY COOLING

Laboratory studies were carried out to evaluate a variety of modes of auxiliary cooling; in all, four approaches were evaluated: 1) Five water-cooled undergarments; 2) an air-cooled vest; 3) an ice packets vest; 4) a wettable cover.

a) EXPERIMENTAL METHOD

All cooling systems were dressed on an electrically heated copper manikin; its skin temperature is controlled by a sensor and proportional controller. A "skin" made out of "T-shirt" material is formfitted to the manikin; this "skin" is left dry for experiments requiring a dry skin condition and completely wetted to provide a 100% wet, maximal sweating, skin condition. All auxiliary cooling systems were worn directly over the manikin "skin" and under a CVC ensemble with a complete charcoal in foam, overgarment chemical protective suit, except that the wettable cover was worn directly on top of a totally impermeable (plastic) chemical protective suit. The electrical power required to maintain constant skin temperature was taken to be equivalent to the heat loss, through the clothing, any other covering items (mask, hood, etc.) and associated trapped and surface still air layers, to the ambient environment.

1. WATER-COOLED UNDERGARMENTS

The five water-cooled undergarments included: a water-cooled cap; a water-cooled vest; the water-cooled cap and vest; short, and long water-cooled undergarments. None provided cooling to the hands and feet. These water-cooled undergarments were worn over the completely wet (maximal sweating) manikin skin. The cooling water flow rate was 22.7 kg/h (378 ml/min) for the water-cooled cap, vest and cap w/vest, and was 63.6 kg/h (1L/min)for the short and long water-cooled undergarments. Cooling water inlet temperatures ranged from 7 to 28°C.

Figure 4a gives the range of cooling provided by each of the five water-cooled undergarments for a completely wet (maximal sweating) skin condition versus the cooling water inlet temperature. The rate of increase in cooling, with decrease in cooling water inlet temperature is: 3.1 w/°C for the watercooled cap; 4.4 w/°C for the water-cooled vest; 7.5 w/°C for the water-cooled cap w/water-cooled vest; 17.6 w/°C for the short, water-cooled undergarment; and 25.8 w/°C for the long, watercooled undergarment. At cooling water inlet temperatures above 10°C (probably too low for "comfort" under most conditions) the water-cooled cap did not provide 100w (86 kcal/hr) of cooling; both the water-cooled vest and the water-cooled cap w/water-cooled vest could provide 100w of cooling. Both water-cooled undergarments (short and long) could provide as much as 400w of cooling. A "comfortable" cooling water inlet temperature of 20°C should provide 46w of cooling using the water-cooled cap; 66w using the watercooled vest; 112w using the water-cooled cap w/water-cooled vest; 264w using the water-cooled short undergarment; and 387w using the long water-cooled undergarment.

The results demonstrate the obvious conclusion that cooling increases with an increase in body surface area covered by a water-cooled undergarment. However, our findings that a) with more skin area covered by a water-cooled undergarment, less area is exposed to receive heat from a hot environment and b) such cooling practically eliminates the effects of adding protective clothing, were not obvious and, indeed, require confirmation with human studies.

2. AIR-COOLED VEST

A hot chamber environment study was initiated using an aircooled vest to distribute cooling air within a CVC suit worn with a complete CW suit. Air flows studied were 6, 8 or 10 ft^3/min and the cooling air inlet temperature to the vest was either 10°C at 20% relative humidity or 21°C at 10% rh. The results are expressed in terms of: 1) the "total heat exchange watts" supplied to the manikin surface; and 2) the "cooling watt" rates. The "total watts" removed from all six manikin sections (head, torso, arms, hands, legs and feet) include both the cooling provided by the cooled air supplied to the air-cooled vest and also the heat exchanges of the total surface area of the manikin with the hot environment. The "cooling watt" rate is the difference between the electrical watts supplied to the torso, arms and legs sections of the manikin when the air-cooled vest is providing cooling to the manikin, and when it is not providing cooling. The experimental data was obtained during exposure to either a hot-humid environment of 29.4°C at 85% rh, or a hot-dry environment of 51.7°C at 25% rh.

The total heat exchanges over the completely sweating surface area of the head, torso, arms, hands, legs and feet when cooling air is supplied to the air-cooled vest are plotted against the cooling air flow rate in Figure 4b, part A; the cooling watts are plotted against the cooling air flow rate in Figure 4b, part B. As expected, both the total heat exchanges and the cooling watts increase with cooling air flow rate and decrease with increasing cooling air inlet temperature.

For an air inlet temperature of 10°C (at 20% relative humidity) and a flow rate of 10 ft^3/min, the total heat exchanges over the manikin surface would be 233w in a 29.4°C (at 85% rh) environment and 180w in a 51.7°C (25% rh) environment. Increasing the cooling air inlet temperature to 21°C (at 10% rh) would reduce the total heat exchanges to 148w and 211w, respectively. Either air inlet temperature easily provides 100 watts of cooling.

(3) ICE PACKETS VEST

The ice packets vest studied holds 72 ice packets; each packet has a surface area of approximately 64 cm and contains about 46 grams of water. These ice packets are secured to the vest by velcro tape. One experiment was conducted with 40% of the ice packets removed. The vest with these ice packets was frozen overnight in a walk-in freezer (air temperature about -20°C) and removed from the freezer about 2 minutes prior to dressing on the manikin. All clothing components dressed on the manikin were originally at the temperature of the chamber air, except for the ice packets vest.

Experimentally, the cooling watts equal the difference in electrical watts supplied when the ice packets vest is providing cooling to the torso and when the unfrozen ice packets vest, at chamber air temperature, is dressed on the manikin. Cooling rates provided (watts) versus time were determined for a completely wet (maximal sweating) skin condition for heat exposure in three hot environments.

Figure 4c shows the decrease in cooling from time 0 minutes, when the ice packets vest was dressed on the manikin, in each of three environments. These decreases in cooling watts with cooling time are based on an average torso temperature of 35°C. The cooling provided by each individual ice packet will vary with time and its contact pressure with the torso surface, plus any heating effect of the clothing and hot environment; the environmental conditions have an effect on both the cooling provided and the duration of time this cooling is provided.

In environments of 29.4°C (at 85% rh) and 35.0°C (at 62% rh), this ice packets vest is still providing some cooling after about four hours of operation. However, in an environment of 51.7°C (at 25% rh), any benefit is negligible after about three hours of operation. When 40% of the ice packets are removed from the vest, the cooling provided over the torso is negligible after two hours of operation. Since the ice packets vest does not provide continuous and regulated cooling over an indefinite time period, exposure to a hot environment would require redressing with backup, frozen vests every 2 to 4 hours when the ice in these packets was completely melted and water temperature approached skin temperature. Replacing an ice packets vest would obviously have to be accomplished when a crewman was in a stand-down position. However, this cooling is supplied noise free and independent of any vehicle energy source or umbilical cord that would limit a crewman's mobility. Its greatest potential appears to be for short duration missions, e.g. aircrewmen on short flights; its drawbacks include the need for a freezer to keep spare vests frozen.

(4) WETTABLE COVER

The wettable cover was simply a two piece cotton cover which extended from just above tops of the combat boots and the wrists to a V-neck at the top. The trouser legs, sleeves and neck opening were generously cut and thus were not in close contact with the totally impermeable, plastic CW protective uniform, which was worn over the combat fatigue uniform.

Predicted values of supplementary cooling, and of the minimal water requirements to maintain the cover wet, for a man wearing the experimental ensemble in various combinations of air temperature, relative humidity and wind speed are given in Figure 4d. A mean skin temperature of 37°C, which would be typical for a stressed man in an impermeable ensemble, has been assumed in these predictions.

V. NON-HEAT STRESS PROBLEMS OF PROTECTIVE CLOTHING

Having presented a variety of options for auxiliary cooling to reduce the heat stress of wearing CW protective garments, if not totally eliminate it under most operational environmental conditions inside (and outside) armored fighting vehicles, it seems appropriate to add that elimination of serious heat stress problems will not totally resolve the degradation in military performance associated with wearing such protective clothing systems. Table 1, an abstract of the operational degradation observed in a series of large scale field studies conducted by the various combat arms in the late 1960s suggests the performance decrements associated with wearing CW protective ensembles in the absence of any heat stress; the majority of these result from mechanical barriers to sensory inputs to the wearer and to barriers for communication between individuals. By redesigning the maneuver scenarios prepared initially by the various combat arms, heat stress was essentially eliminated, and it became feasible to assess other forms of performance decrements. The table compares the performance of troops wearing 1) normal combat clothing and equipment (MOP I), 2) CW protective ensembles "open", i.e. without hood, gloves and with all apertures open, but with gas mask (MOP II) or 3) fully encapsulated (MOP III) with mask, hood and gloves, and all uniform openings sealed, for four critical elements of combat: 1) fire power, 2) communications, 3) mobility and 4) support. There is a great deal of variability in the results of any such "large scale maneuver" field studies of operational performance, and some of the expected "overcompensation" can be noted; i.e. performance is actually improved slightly by imposing impediments that the troops are aware of and can make adjustment for. However, overall it is clear that elimination of heat stress, while it will allow mission performance to continue, will not totally eliminate the constraints imposed by CW protective clothing systems.

REFERENCES

Goldman, R.F. Tolerance time for work in the heat when wearing CBR protective clothing. Mil. Med. 128:776-786, 1963.

Joy, R.J.T. and R.F. Goldman. A method of relating physiology and military performance: A study of some effects of vapor barrier clothing in hot climate. Mil. Med. 133:458-470, 1968.

Martin, H de V. and R.F. Goldman. Comparison of physical, physiological methods of evaluating the thermal stress associated with wearing protective clothing. Ergonomics 15:337-342, 1972.

Breckenridge, J.R. and R.F. Goldman. Effect of clothing on bodily resistance against meterological stimuli. Progress in Biometeorology, Vol. 1, Part III, Section 19, 194-208, Chap. 7, 1977.

Shapiro, Y., K.B. Pandolf, M.N. Sawka, M.M. Toner, F.R. Winsmann and R.F. Goldman. Auxiliary Cooling: Comparison of air-cooled versus water-cooled vest in hot-dry and hot-wet environments. Aviat. Space and Environ. Med. In press, Apr/May 1982.

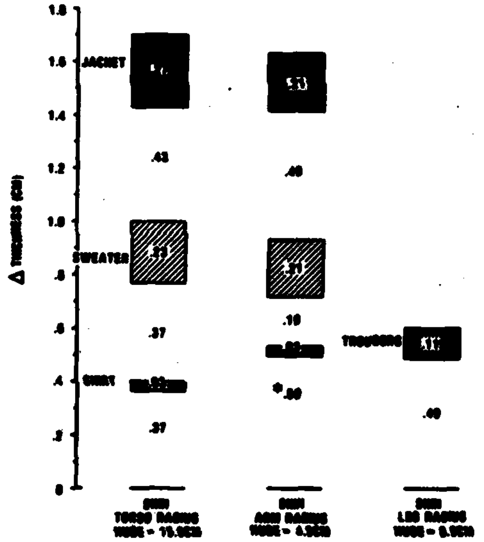
The views, opinions, and/or findings contained in this report are those of the author(s) and should not be construed as an official department of the Army position, policy, or decision, unless so designated by other official documentation.

Human subjects paticipated in these studies after giving their free and informed voluntary consent. Investigators adhered to AR 70-25 and USAMRDC Regulation 70-25 on Use of Volunteers in Research.

Table 1. OPERATIONAL DEGREDATION WITH CW PROTECTION WITHOUT HEAT STRESS

MOP	FIRE POWER I	II	III	COMMUNICATIONS	I	II	III	MOBILITY	I	II	III	SUPPORT	I	II	III
INFANTRY															
Rounds per hit				Messages unanswered	6	-	10	Road march time	-	-	12%↑	Time to get medic			270%↑
M-16	3.4	3.9	8.6	Time to install wires	15'	23'	26 min	March rate	-	< 6%	10%↓	M-16 maintenance			10%↑
(%)	-	9↑	59↑	Getting specific man	1-2'	-	6+ min	Assault time	-	< 10%	>100%↑	CBR recon	6 min	-	9.5 min
M-60	10	5.7	10.7	Voice versus hand signals	-	-	50%↑	platoon leader			142%↑	Ammo resupply	13 min	30%	67%↑
(%)	-	43	7↑					first man			197%↑				
M-79	5.3	3.0	8.6					last man			102%↑				
(%)	-	43%	20↑												
ARMOR: N.B.- MOP III also required closed hatch															
Rounds per hit				Target identification # missed	1	7	23	Road march time			9%↑	Boresight			1 min↑
M-73	-	NSD	-		13%	20%	62%	March rate			13%↓	**Complaints**			
M-85	-	0%↑	100%↑	# of transmissions	-	19↑	20↑	Attack time			20%↑	Gloves	33% maintenance 27% loading		
105	-	0.3r↑	0.6r↑	duration of transmissions	-	12%↑	10%↑	Attack difficulty			3X	Mask/Hood	30% vision		
Zero main gun	-	1 min↑	3 min↑									CBR suit	19% movement		
ARTILLERY:															
Time from receipt at FDC to battery order				In firing sections responses	26 OK	11%↓	20%↓	Last unit across SP			192%↑	Filling sandbags	22 min	25'	35'
auto adjust.	-	27%↑	137%↑	Accuracy	>95%	-	65%	Clear hasty position			65%↑	Wire splicing	3.5 min	7.75'	8'
registration	-	100%↑	75%↑					Enter-to all ready			31%↑				
transfer	-	33%↑	96%↑												
Target location F.O. i.d. to call for fire	90 sec	130 sec	183 sec	F.O. i.d. to end of mission	27 min	-	60 min					NETOOSE Data Unclassified 5 Jan 81			

ENGINEERS: road repair, bridge building, demolition -NSD; voice commands repeated 2X as often in MOP III.



*THICKNESS OF TRAPPED AIR FILM > 0.5CM IS INEFFECTIVE BECAUSE OF INTERNAL CONVECTION

FIG. 1 CLOTHING THICKNESS AS A FUNCTION OF FABRIC AND TRAPPED AIR LAYERS.

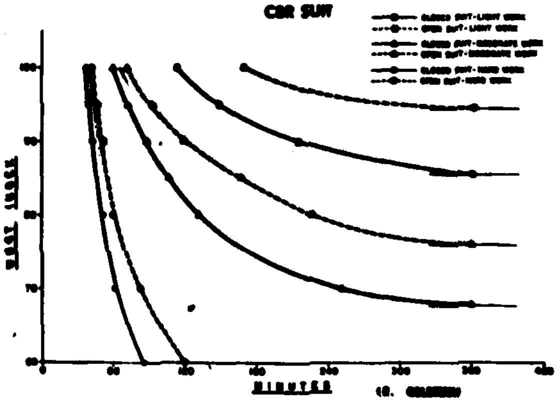


FIG. 2 PREDICTED TIME TO 50% UNIT HEAT CASUALTIES.

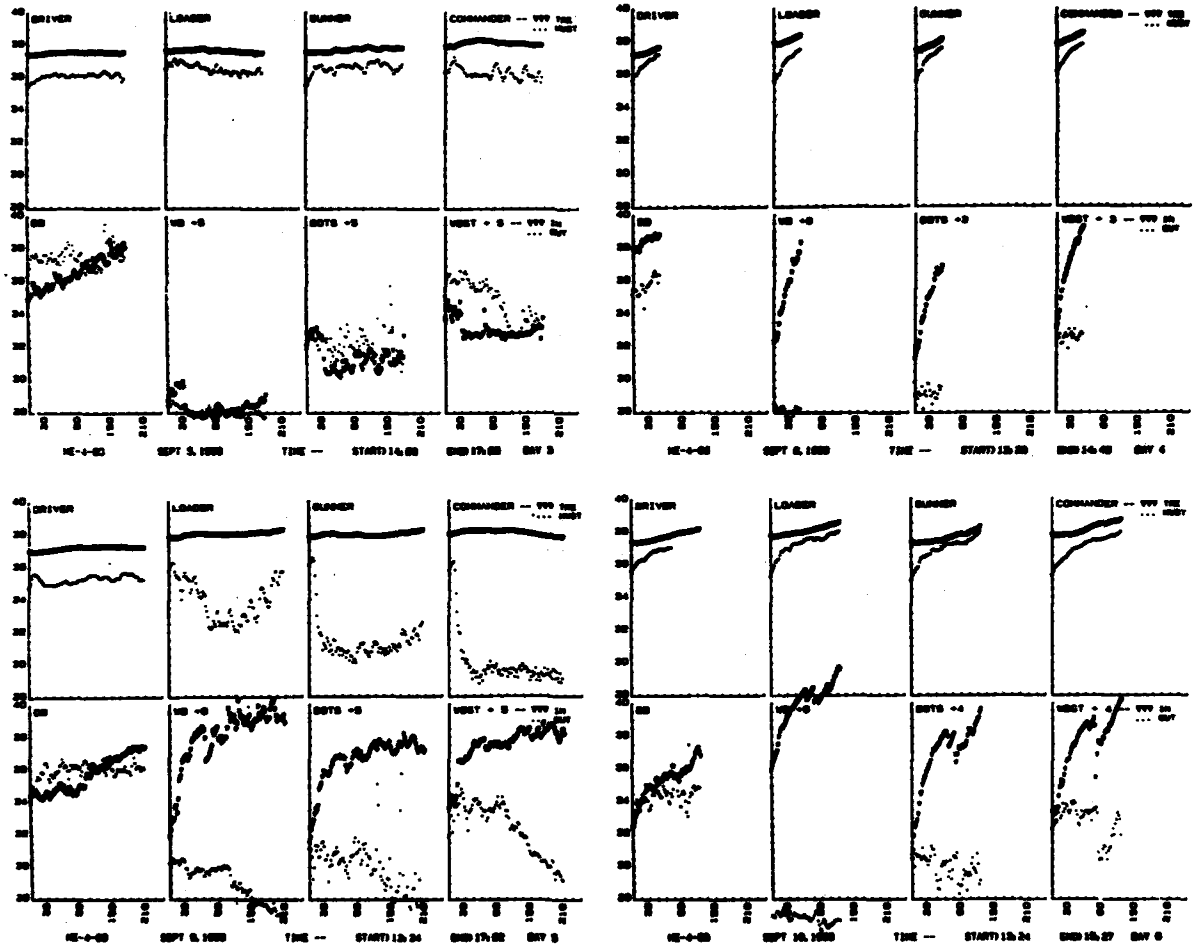
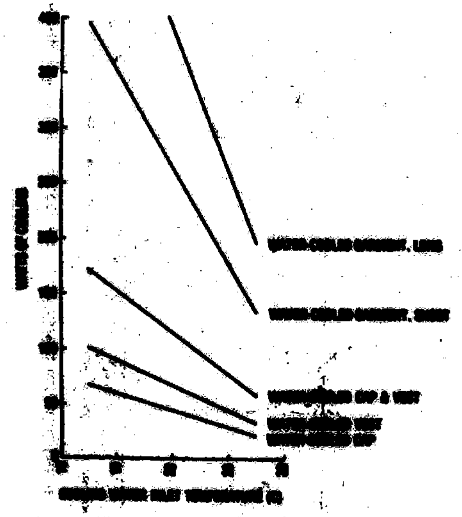
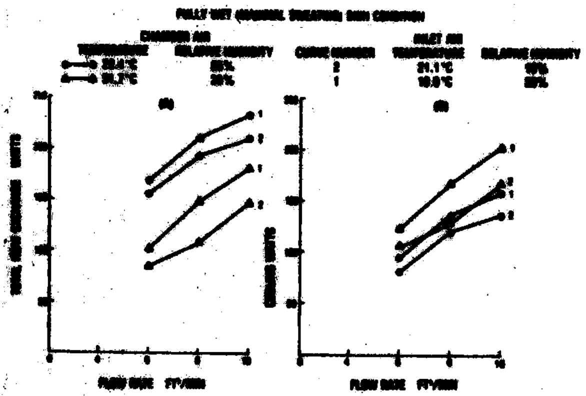


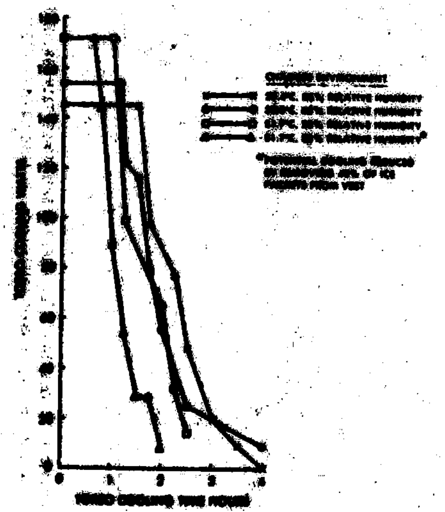
FIG. 3 INDIVIDUAL CREWMAN DEEP BODY (TRE) AND SKIN (MWST) TEMPERATURES, AIR (DB) AND WET BULB (WB) TEMPERATURES AND THE TWO HEAT STRESS INDICES (WBGT AND BOTS BALL) INSIDE AND OUTSIDE THE TANK ON DAYS 3 TO 6.



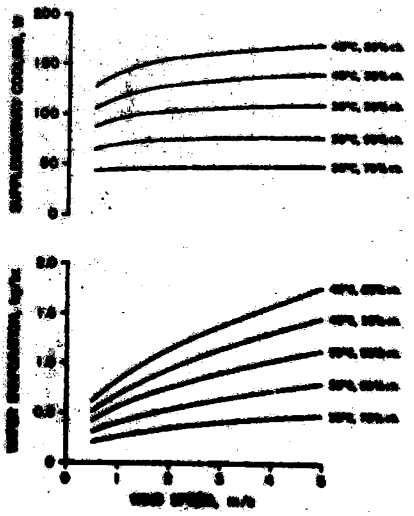
[illegible]



[illegible]



[illegible]


200
150
100
50
0
2.0
1.5
1.0
0.5
0
0
1
2
3
4
5

[illegible]

DEVELOPMENT OF METHODS FOR ASSESSMENT OF GLIDING PARACHUTE APPLICATIONS

THOMAS F. GOODRICK
US ARMY NATICK R&D LABORATORIES
NATICK, MASSACHUSETTS 01760

In order to assess the possible utility of gliding parachutes, it was necessary to develop a set of methods for predicting behavior of gliding parachutes. Gliding parachutes presented some unique problems not covered by the technology base existing in 1968. At that time a few brave jumpers were experimenting with squares. The Army was encountering difficulty in obtaining consistent performance from a small gliding system for cargo. On some days it worked; on others it did not. Engineers with parachute experience viewed the problem mainly as inconsistent deployment. Engineers with aircraft experience thought the problem was poor navigation and lack of power. Neither field of engineering was wholly appropriate.

Based on the observations that canopy opening was not too different from conventional parachutes in most critical respects and that canopies tended to glide with reasonable steadiness, we chose to concentrate first on the problem of guidance. As a first step, we worked out a graphical (manual) method of plotting trajectory points for a gliding canopy drifting with the wind while homing on a target. A visiting student helped with the manual labor. Soon we transferred the technique to a programmed calculator though graphs still had to be drawn by hand. The technique was crude, but it effectively demonstrated some key features of radial homing as applied to low-speed gliding parachutes.(1) This was our first use of "computer" simulation. Later we moved to a fancier programmable calculator with an X-Y plotter and studied more exotic guidance schemes.(2)(3)

However, as we investigated guidance schemes it became obvious that we needed more knowledge of the flight dynamics of gliding parachutes. Published test data was inadequate. A test program appeared to be too expensive and time-consuming. Also, we were not sure what parameters should be measured in testing - nor how to measure them. Thus, we began the development of analytical and computational techniques for study of all aspects of gliding canopies. This development effort has included models for study of longitudinal stability and for study of unrestricted flight in six-degrees-of-freedom (6DOF) which have been verified recently by experiments.

Stability Analysis

The longitudinal stability is the key aspect determining glide performance of gliding parachute systems.(5) A balance must be achieved between the various moments so that the parachute maintains a steady glide at an angle of attack permitting useful glide velocities and returns to this angle if disturbed. The factors that determine the condition of balance or stability for a parachute are:

1. The orientation and location of lift and drag forces acting on the canopy;
2. The magnitude of material mass and its distribution (including payload mass);
3. The magnitude of entrained mass within the canopy which is offset by buoyancy in steady flight but must be included in determining the location of the system center of mass;
4. The location, magnitude and orientation of payload drag.

The parameters available which can be adjusted to achieve stability or "trim" at a desired glide condition differ significantly from those available in aircraft design. The key parameter is the aerodynamic force on the canopy. For the Parafoil canopy (Fig 1), wind tunnel data from nearly full scale canopies has been published(4) showing the variation of lift and drag components of aerodynamic force with angle of attack. A pair of empirical curves was fitted to the data to allow smooth calculation of forces at any angle of attack (Fig 2). Combined with an analysis of geometry and mass

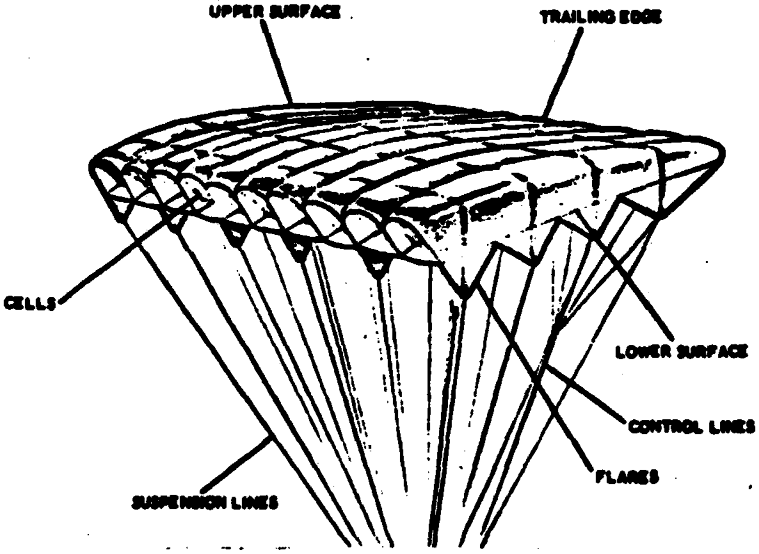


Fig 1 Parafoil Canopy

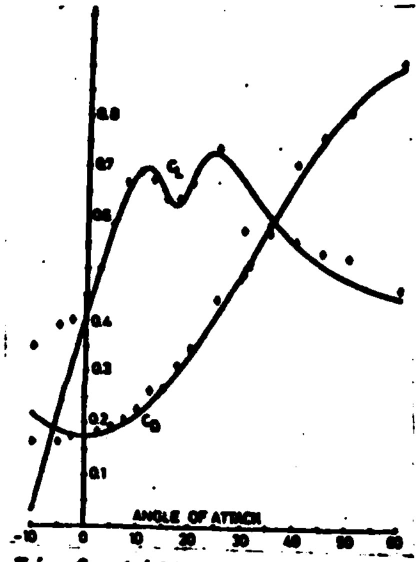


Fig 2 Lift and Drag Coefficients

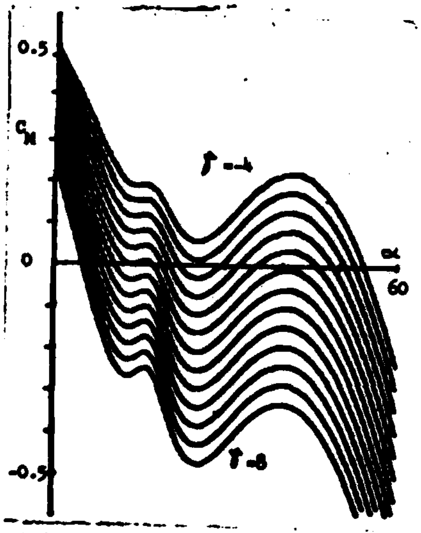


Fig 3 Static Stability (No Flap)

distribution, this permits determination of stable conditions with variation of control parameters.

There are two control parameters which may be varied independently or in combination to achieve stability. One is the angle of incidence (γ) controlled by basic suspension geometry holding the chord plane of the canopy at an angle to the axis connecting the canopy and payload centers of mass. It is convenient to define this angle as the inclination of the chordline below a line perpendicular to the axis passing between the centers of mass. For a particular setting of incidence angle, longitudinal static stability is indicated by passage of the moment coefficient (C_M) through zero with a negative slope against angle of attack (Fig 3). The second control parameter is the flap or trailing edge deflection which alters the lift and drag forces produced at any given angle of attack. Comparison of the moment coefficient variation in the two cases of zero and full flap (Fig 4) shows how variations of either or both of these parameters can affect the value of the stable angle of attack. Sport canopies are normally designed with set angles of incidence and with flap deflection varied in flight to adjust the steepness of the glide. We have tested canopies designed with variable suspension geometries for in-flight variations of angle of incidence.

The steady-state glide velocities attainable at the various stable angles of attack (following some oscillation) show that a wide range of performance is available using variable incidence (Fig 5 and 6). Optimum performance for a given canopy/payload combination can be obtained by adjustment of both control parameters.

Distance separating the payload and canopy mainly affects the steep-

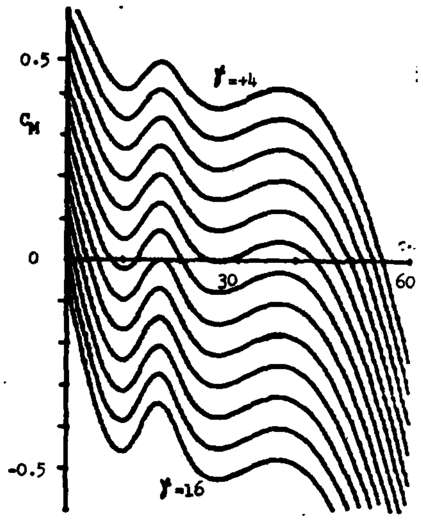


Fig 4 Static Stability (Full Flap)

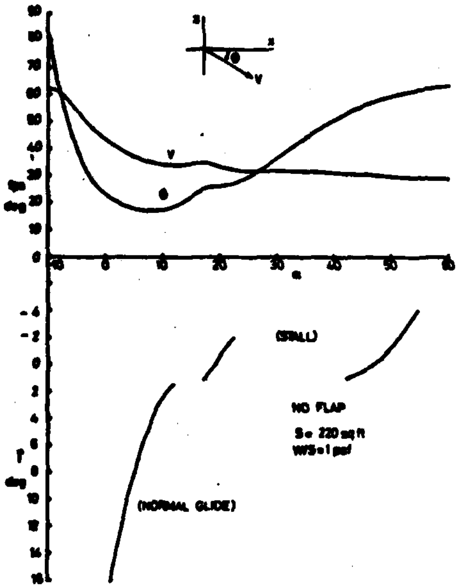


Fig 5 Range of Equilibrium Velocity (No Flap)

ness of the slope of the pitching moment with little affect on the trim angle of attack. Payload drag has a moderately significant effect. In this case a partially streamlined load is assumed.

The stability analysis has been used throughout the flight simulation studies including the 6DOF studies discussed below and has been verified in the experimental studies below. In addition to indicating the parameter settings needed to attain any desired glide performance, the stability analysis has indicated two potential problem areas requiring further study in dynamic simulations and flight tests.

1. Though larger systems have the same wing loading as small systems in order to retain similar glide velocities, canopy air mass varies as the cube of the linear scale. Thus large systems will have different static and dynamic stability which must be considered in their design.

2. Unlike aircraft which start their flight under well-controlled conditions, gliding parachutes start flying at conditions differing greatly from their design trim conditions. Hence, undesirable bistable conditions may be encountered inhibiting attainment of the desired steady glide condition unless parameters are carefully set or some means of adjustment is provided.

6DOF Simulation

A six-degree-of-freedom (6DOF) flight simulation program has evolved out of earlier, more limited simulation studies and theoretical analyses. The key features and theoretical bases of the program are discussed below. The studies using the program have yielded new information as to the nature of dynamic stability, turn control effectiveness, and effectiveness of guidance schemes. The six degrees of freedom are the three translations and the three rotations of the system of two "rigidly connected" bodies.

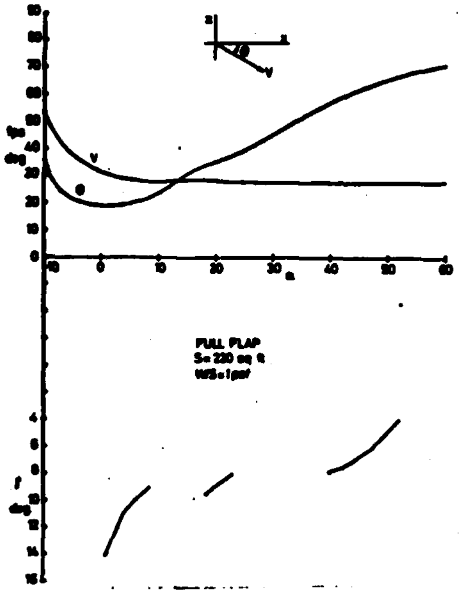


Fig 6 Range of Equilibrium Velocity (Full Flap)

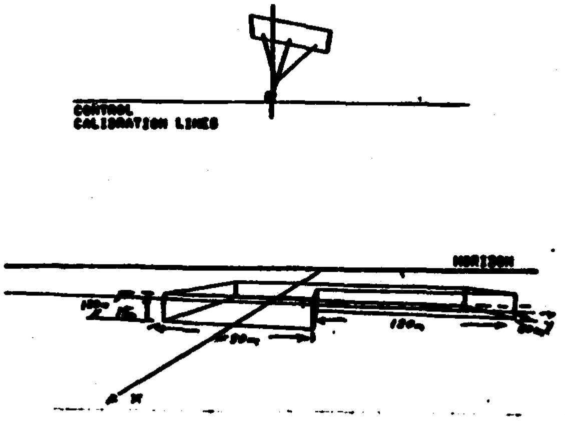


Fig 7A Simulation View from Fixed Point

Computer Program

The 6DOF simulation program for gliding parachutes features scaling functions for geometry and mass distribution, allows selection of payloads and associated drag areas, selection of variable and gusting two-dimensional wind functions, presents a deployment window plot for selection of initial conditions, features several pictorial display options during flight computation, includes a variety of manual and automatic control options, and presents plots of the parameter histories in a flexible format. A recent modification allows simulation of random errors on measured parameters used in guidance computations.

Applications of the program include study of flight mechanics and study of guidance techniques. For example, to study the turn control response of a particular canopy/payload system, calm air is selected, manual control mode is selected and a pictorial viewpoint fixed in space near the expected flight path is established. Generally, a view is selected showing a line drawing of canopy and payload relative to an inertial background (Fig 7). The display can be either a centered single image against a moving background or of multiple frozen images passing across a fixed background. The preselected type of control deflection is input via cross hairs against a calibration scale, with inputs possible at preselected intervals during the flight. The rate of change of the control deflection is controlled by a time-dependent function. The operator observes the progress, sets desired control deflections and changes the view point or format as needed. Hard copies of the display are made as required. Values of key data parameters are also displayed with the pictorial view in one of several formats. When the

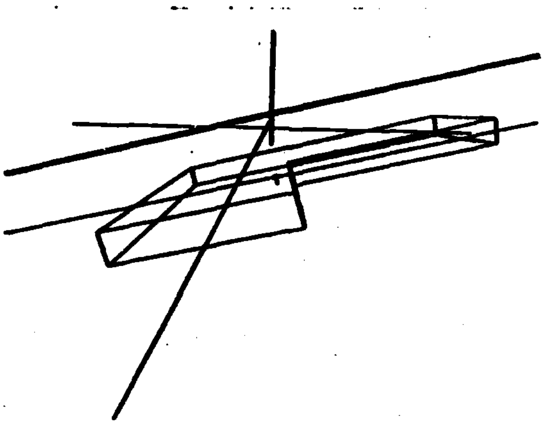


Fig 7B Simulation View from On-Board

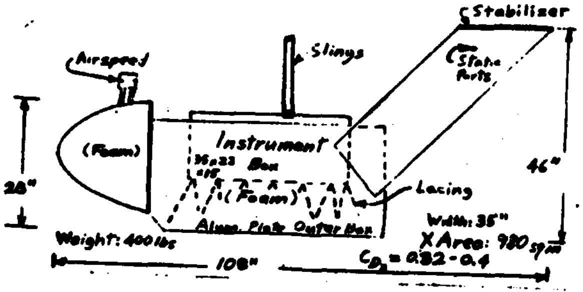


Fig 8 Flight Data Unit

maneuver is completed the operator selects the output mode in which a ground track and histories of any of 31 key parameters are plotted. The operator selects certain combinations of parameters and sets time intervals per plot.

For simulation of manual control, either an air-to-ground view from on-board the payload or a ground-to-air view from the landing zone can be selected. A variety of three-dimensional inertial references can be established. Limited data - perhaps only elapsed time - is displayed during the flight with full data available at completion. For simulation of automatic guidance, the appropriate scheme is selected and initialized if necessary. Return to manual control will occur either at a selected time or altitude. A running pictorial display is provided to monitor computer performance without delaying the computation.

Geometric and Aerodynamic Assumptions

The basic components of the program are the assumed aerodynamic and geometric characteristics and the kinematic relations. The canopy/payload system is modeled as two points connected by an inextensible line subject to tensile forces (if compressive forces occur, the payload freefalls until the orientation is such as to allow tensile forces). The payload mass contributes to system moments of inertia about the longitudinal and lateral axes only since suspension geometry usually decouples the motion about the vertical axis. A constant drag area is presently assigned to the payload although studies of particular payload configurations with lift, drag, and side forces are anticipated. The mass of the canopy is modeled as a simple rectangle inclosing an air mass. For apparent mass, up to one-half the displaced air mass (volume of air displaced by the canopy and its captive internal air) is selected as a constant added mass concentrated at the centroid in calculations of moments of inertia. The geometric distribution of the captive included air mass is considered in calculation of moments of inertia. This mass and the apparent mass change with altitude according to ambient density.

Certain unique assumptions have been made in the treatment of aerodynamic forces and moments. Empirical functions were developed for the basic wing lift and drag coefficients of Parafoil canopies measured in the wind tunnel. A side force coefficient varying with the sine of the side slip angle was also extracted from the wind tunnel data. Early attempts to incorporate these coefficients and other moment coefficients and stability derivatives measured in the wind tunnel under static conditions failed to give realistic results. Extremely low turn rates were noted under some typical quasi-steady conditions even though the data supposedly reflected the conditions of control deflection. Realistic performance - as compared with motion pictures of jumpers and limited available test data - was finally achieved by treating the canopy as geometrically distributed segments including spanwise curvature with local lift and drag computed for each segment based on local velocity and angle of attack. Local lift coefficients vary from the total wing lift coefficients in a manner approximating an elliptical distribution. The empirical functions for total wing lift and drag coefficients include effects of trailing edge deflections. To study turn and pitch (or "braking") control with trailing edge deflection, the deflection is applied via the empirical functions at selected segments. It is important to note that, due to the spanwise curvature, resultant aerodynamic forces on deflected segments may be significantly inclined to the vertical-longitudinal plane.

A study using a separate computer program was made of the quasi-static stability effects of typical velocity and angle of attack distributions resulting from the segmentation of the wing. It was found that for typical wing geometries associated with payloads from 200 to 2000 pounds, the effects of velocity distribution were highly stabilizing.(6)

Experimental Verification of Simulation

To verify the simulation program a flight test was performed using a Parafoil canopy of 200 sq ft (18.6 sq m) and a payload of 380 lbs (165 Kg) specially designed to measure those key parameters needed for verification. The emphasis of the test program was on dynamic flight conditions associated with turning maneuvers. The payload included remote radio control hardware manually activated from the ground.

Flight Data Unit

The flight data unit (FDU) serving as the payload consists of an inner box cushioned within an outer shell to which a specially shaped foam nose cone and stabilizing tail fins are attached (Fig 8). The geometry of the fins and nose cone was determined during wind tunnel tests to achieve good static and dynamic stability. For accurate airspeed and gyro data the payload must track accurately with the canopy. The payload must also track well when carried by sling below a helicopter.

A J-Tec vortex-sensing airspeed indicator is mounted on the upper part of the nose cone. Wind tunnel tests were made for calibration at various angles of attack. Twin static ports in the surface of the vertical fins provide static pressure for the Rosemount 1241 altimeter with rate output. This mounting arrangement is not affected by moderate slideslip. The brand names of these units are mentioned because they appear to be the only units available that are adequate for this application. This was determined in several independent surveys. The J-Tec airspeed indicator acoustically senses the rate of vortex shedding from a special cross-piece in the channel of the indicator. It is highly accurate and responsive in the test range of 10 to 18 m/s. Pressure transducers would be inaccurate in this range. The Rosemount altimeter is designed for test applications with good accuracy both in altitude and electronically-derived rate measurement. Airspeed and descent rate are the two main parameters of gliding flight. Also, the horizontal component of airspeed can be calculated from the vector difference between total airspeed and descent rate only if each parameter is measured accurately.

The FDU also contains three rate gyros on the yaw, pitch and roll axes. An integrating circuit is attached to the yaw rate gyro to provide approximate heading data. The gyro data is useful in indicating the nature of turn response which can vary widely depending upon the mechanics of turn control.

Control deflection is measured from the servo potentiometers. Originally, current draw of the servo motors was measured as an indication of control force. This has been found unreliable as an indication of force because of electrical transients. Modifications are being made to measure forces from transducers linked to the lines.

A 14-channel FM tape recorder is used to record all signals. A playback unit re-transforms the FM signal to voltage analog to drive a strip-chart recorder for final copy. The tape recorder is satisfactory for recording data during a flight including extreme oscillations. However, it is not satisfactory for recording data under high-shock conditions such as during deployment and landing. Apparently, the tape is jiggled as it passes the head when the unit is jolted.

The FDU is controlled by a model aircraft radio-control system. The radio-controlled servos drive potentiometers which control the large motors drawing in control lines. Currently, two channels are used for turn and glide ratio control. Three motors and lines are used for the left turn, right turn and glide ratio. The Parafoil is rigged for deflection of the third (or "C") flare on the second set in from each tip or turn. This produces primarily a roll tilt of the wing with a slight incidence shift. The trailing edge is separately deflected for glide control.

Weight of the payload was 355 pounds during initial testing. Later modifications added 40 pounds. This weight is carried well by the 200-square-foot Parafoil with unreefed deployment at low speeds (10-30 knots). Landing speed is somewhat high with this loading.

Except for the airspeed indicator, all equipment is housed within the inner box of the FDU (35 X 22 X 14 inches). The inner box is tied to the outer shell at the bottom and surrounded by several inches of polyurethane foam. The outer shell is 108 inches overall, including nose and tail with a cross-section of 35 X 28 inches. Drag area (C_DS) measured approximately in the wind tunnel is 2.15 - 2.69 square feet (0.2 - 0.25 sq m).

Setting System Parameters

Those aspects of the simulation model needing verification included: (1) the lift and drag functions, (2) the model resulting trim condition predicted by the longitudinal stability model, (3) the assumption that linearized moment coefficient and deviations could be replaced by local force computations at spanwise locations, and (4) the computational techniques employed. Thus, in matching flight data none of these factors were altered. Input parameters altered for the match, as listed below, were those parameters not physically measurable. Apparent mass (sometimes called "virtual" mass) is a flow field phenomenon - not a physical mass. System length depends on in-flight mass distribution and distorted suspension and canopy geometry. While these parameters could be "ballparked", they could not be physically measured. Ideally, to examine flight dynamics it is desirable to start from a steady condition. For gliding parachutes this is impossible. After over twenty flights, steady conditions have never been observed. Repeatedly, a maneuver must be entered before the efforts of a previous maneuver have dissipated even though flight appears steady from the ground. Wind turbulence adds additional disruption. The two flights matched below are the best illustrations for verification of the simulation.

To obtain a reasonable match, the following system parameters were varied in the simulation:

1. Apparent mass coefficient (one corresponds to half the volumetric displacement) was initially assumed to be 1.0. Values of 0, 0.8, 1.0, 1.5, 2.0, 3.0 were tried. Final value is 1.0!

2. Payload drag area (C_DS) was initially estimated at 0.4 sq m. Values of 0.2, 0.25, 0.3, 0.35, 0.4 were tried. Value selected is .2.

Wind tunnel tests later confirmed the value to be 0.20 to 0.25 sq m.

3. System length, which is the distance from payload center of mass to canopy center of mass, was varied between 5.5 m and 6.5 m. Final value is 6.5 m, which corresponds with physical measurement.

4. Trailing edge deflection was assumed and confirmed to be zero, though deflection of the edge near the tip was assumed in some trails. This method of turn control creates high initial yaw rates which were not observed in the flight data.

5. Roll-tilt turn deflection was found to be 5.0 deg in order to obtain a match. Values of 4.0 to 6.0 were tested. This parameter cannot be measured since it is the tilt angle of the resultant force vector. The flight data indicated a moderate left turn at neutral control line deflection which was matched by assuming -1.5 deg (left) tilt at neutral. Thus, in the final simulation presented here, tilt varies from -1.5 to 5.0 deg for 100% deflection.

6. Incidence coupling was found to occur with tilt-turn deflection. This was the main surprise learned in the matching process. The incidence angle is the angle of the chord below the canopy X-axis with the Z-axis passing from center of mass to payload center of mass. In the stability analysis of Reference 3, the optimum value was found to be 4 deg which appears close to the constructed value. Film of the flight test showed substantial incidence change of the section to which the control line attached. Also, it was impossible to match the characteristic initial decrease in descent rate as turn deflection was applied if incidence angle was held constant at 4 deg. But when incidence was allowed to decrease to 2 deg linearly as turn deflection went to 100%, the proper decrease in descent rate was obtained, and the proper sensitivity toward spin was observed. Though no spin occurred in the first flight, considerable sensitivity to duration of the applied stroke was observed. Thus, to obtain the data match, incidence was assumed set at 4 deg (nose down) varying to 2 deg (nose down) with 100% turn deflection.

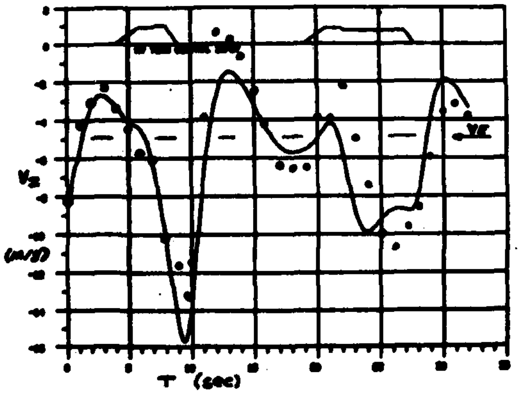


Fig 9 Descent Rate (Exp/Sim)

Comparison of Data

The descent rate data matches well in parts with certain deviations (Fig 9). Descent rate is the key parameter for comparison since all force, moment and orientation variations have a direct effect on it. Note the large deviation from equilibrium ($V_\infty = -4.7$ m/s). This indicates that the turns are highly unsteady due to excessive deflection. Turns of moderate rate would show excursions to -6 or -7 m/s after deflection with little sensitivity to duration of the deflection. In the simulation just previous to the one presented, the second turn deflection was a few percent greater initially, producing a peak descent rate of -16 m/s. If the first turn were held one to two seconds longer, a high rate spin would occur with descent rate peaking at -22 m/s and a turn rate of 150 deg/sec. A spin almost that bad occurred on a later flight.

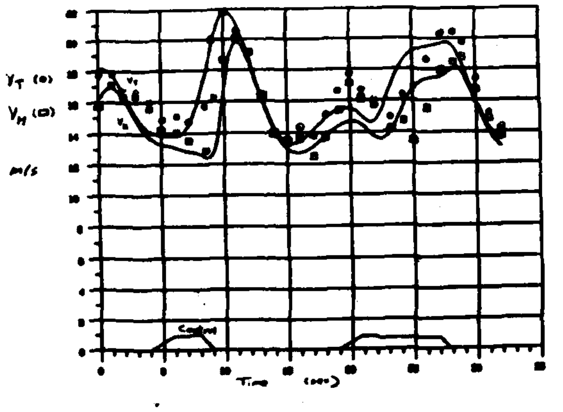


Fig 10 Total and Horizontal Airspeed

Fig 10 shows the total airspeed measured and the horizontal airspeed computed as the vector difference of total with descent rate. It was difficult to match initial conditions since the system was just recovering from a previous right turn. It is encouraging that horizontal airspeed can be accurately estimated from measurement of total airspeed and descent rate because this allows evaluation of glide performance independent of wind and also allows estimation of wind speed when compared to measured speed by a guidance system (not done for these flights).

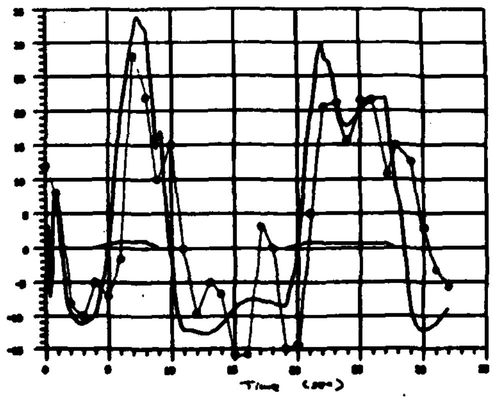


Fig 11 Yaw Rate

Fig 11 shows the axial yaw rate data. Fig 12 shows yaw, pitch and roll rates for the first turn. The gyro data includes significant "noise" caused by payload oscillations. The payload oscillations are of high rate (10 deg/sec) but of short period so that system motion can be identified. It was not known whether gyro data would be meaningful or not. For turning motions encountered in this flight there is reasonable correlation at the peaks on all axes.

Spin Performance

An unintended spin produced data which was also matched closely by simulation. During deployment, a suspension line wrapped over one cell at the tip, distorting the outer 25% of the canopy. Measured descent rate averaged 9.7 m/s with oscillations of between 0.5 and 3 m/s with a period

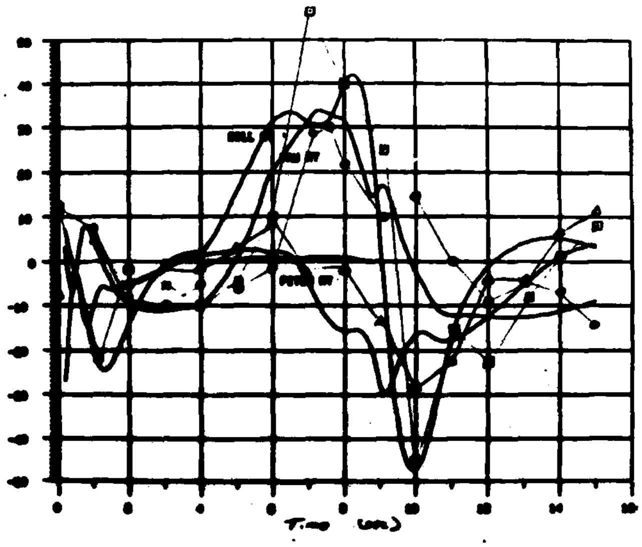


Fig 12 Rates of Yaw (○), Pitch (△) and Roll (□)

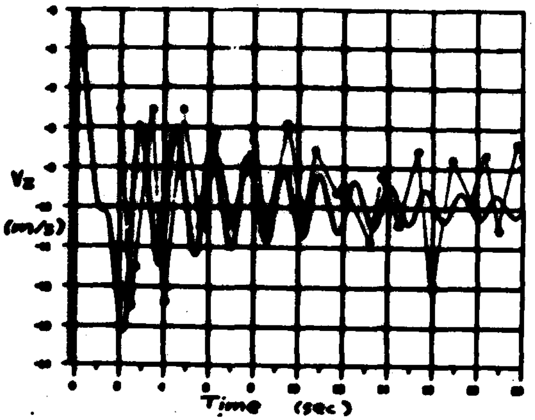


Fig 13 Descent Rate During Spin

of 1.6 seconds. Film indicated that the trailing edge was partially deflected as a result of twisted lines. Otherwise, most of the canopy was inflated to a proper shape. Simulations were made using zero, half, and full flap. Descent rates averaged 22, 10, and 8 m/s, respectively. At half flap the simulation also showed (Fig 13) an oscillation of 2 m/s, about the average of 10 m/s damping to 0.5 m/s in 11 seconds. Period of oscillation was 1.6 seconds. The Euler (or inertial) yaw rate during the flight was estimated from the video tape at 197 deg/sec (averaged over ten revolutions). Simulation at half flap indicated 205 deg/sec. Other data normally recorded was off-scale during this flight. The close match of rate of descent and yaw rate adds additional validity to the simulation model. The fact that this occurred in spite of the canopy distortion may be explained by the assumption that the distorted portion would be stalled anyway in a spin, producing only drag. Spin dynamics appear to be driven mainly by the mass distribution. The flight data did not exhibit the same damping characteristics; however, the system may have been subject to various excitations as the payload pitched, wind varied, or other changes occurred. The simulation was controlled by holding full left turn and half flap while starting at a nominal, zero-flap glide condition. The spin was established within 3 seconds. Curiously, the ground track showed tight left-turning circles whose centers precess circularly to the right.

Comparison has been made in a few other cases with results similar to those presented here. Therefore, the experimental results appear to validate the simulation program as applied to a 200-square foot Parafoil of aspect ratio of 2.0 with a 350-pound payload. This conclusion is warranted because:

1. Due to the discrete character of control inputs, response is driven mainly by aerodynamic and internal properties assumed for the model.
2. The combination of multi-axial responses predicted by the model is in close agreement with experimental results.

As of this writing, about 10 good data flights have been made. The simulation model has matched salient features of most flights. There is some difficulty when wind changes occur. The data shows a response to a disturbance rather than a control input without indicating sufficient data for simulation. This has made it difficult to match flare dynamics since the few flare landings made were accompanied by some wind turbulence. The simulation has been used to identify some anomalies such as incidence shifts due to line stretch and trim change caused by a pilot chute.

Conclusion

The extent of agreement between simulation and flight data indicates that the primary factors included in the longitudinal stability analysis and in the 6DOF simulation are correct. Although the stability analysis predicts only steady-state behavior, it forms the basis required for analysis of dynamic behavior in the body-fixed XZ plane. The agreement seen in descent rate (Fig 9) would not be possible with an invalid stability model. Of course, in turning flight during the response immediately following the deflection, other factors such as the assumed spanwise distribution of lift and drag become predominant. The agreement in yaw rate (Fig 11) best illustrates correctness of this aspect of the 6DOF model. The mechanics driving motion during a spin are quite difficult to understand; however, the agreement shown in the descent rate (Fig 13) indicates that the mass ratios assumed are accurate and further justifies the assumption of spanwise distribution of lift and drag.

In further development activities on gliding parachute systems, the 6DOF model will serve to guide exploratory work and will be updated for more accurate application to different canopies and to larger systems.

References

1. Goodrick, T., "Wind Effects on Gliding Parachute Systems with Non-proportional Automatic Homing Control", TR70-28-AD, USA Natick R&D Laboratories, 1969.
2. Goodrick, T., "Estimation of Wind Effect on Gliding Parachute Cargo Systems Using Computer Simulation", Paper #70-1193, American Institute for Aeronautics and Astronautics (AIAA), 1970.

3. Goodrick, T., Pearson, A., Murphy, A., "Analysis of Various Automatic Homing Techniques for Gliding Airdrop Systems with Comparative Performance in Adverse Winds", Paper #73-462, American Institute for Aeronautics and Astronautics (AIAA), 1973.

4. Ware, George M. and Hassell, Jones L. Jr., "Wind Tunnel Investigation of Ram-Air-Inflated All-Flexible Wings of Aspect Ratios 1.0 to 3.0", NASA TM SX-1923, Langly Research Center, 1969.

5. Goodrick, T., "Theoretical Study of the Longitudinal Stability of High-Performance Gliding Airdrop Systems", Paper #75-1394, American Institute for Aeronautics and Astronautics (AIAA), 1973.

6. Goodrick, T., "Simulation Studies of the Flight Dynamics of Gliding Parachute Systems", Paper #79-0417, American Institute for Aeronautics and Astronautics (AIAA), 1979.

ANALYTICAL AND EXPERIMENTAL STUDIES OF THE RESPONSE OF A CYLINDER TO NUCLEAR THERMAL/BLAST LOADS (U)

*FREDERICK H. GREGORY, MR.
RICHARD J. PEARSON, MR.
U.S. ARMY BALLISTIC RESEARCH LABORATORY
ABERDEEN PROVING GROUND, MARYLAND 21005

Two phenomena which are emitted from a nuclear weapon, whose target vulnerability ranges often overlap, are thermal radiation and blast waves. In most situations of interest to the Army, targets are exposed to a high percentage of the total available thermal pulse prior to blast wave arrival. This sequence of loads is detrimental to the survival of light weight structures in which significant in-depth heating occurs prior to the blast wave envelopment. There has been evidence that this thermal preheating causes enhanced structural damage [1,2].

EXPERIMENT DESCRIPTION

In order to demonstrate these types of synergistic responses and to provide some quantitative data, three experiments have been conducted on hollow cylinders of aluminum 6061-T6 in the BRL Thermal/Blast Simulator. One experiment was conducted with thermal-only exposure, one with blast-only loading, and one with combined thermal/blast loading. A diagram of the Simulator is shown in Figure 1. The target cylinder is clamped at each end and placed with its axis in a vertical position as shown. The thermal source is situated downstream from the target to minimize the effects of the shock passing through heated air and combustion products. In the combined thermal/blast test, the target is rotated through 180° so that the heated side of the target cylinder is exposed to the on-coming shock wave.

The thermal heating of the aluminum cylinder is accomplished by burning of a mixture of aluminum powder and oxygen. Two thin plastic membranes serve to contain the oxygen gas before ignition. The aluminum powder is sprayed through nozzles into the oxygen atmosphere and ignited starting with the nozzles farthest from the target. The aluminum oxide particles resulting from the burn irradiated the near side of the cylinder, raising its temperature by approximately 200°C.

The shock loading on the cylinder is produced by pressurizing the driver section of the shock tube shown in Figure 1 and then explosively

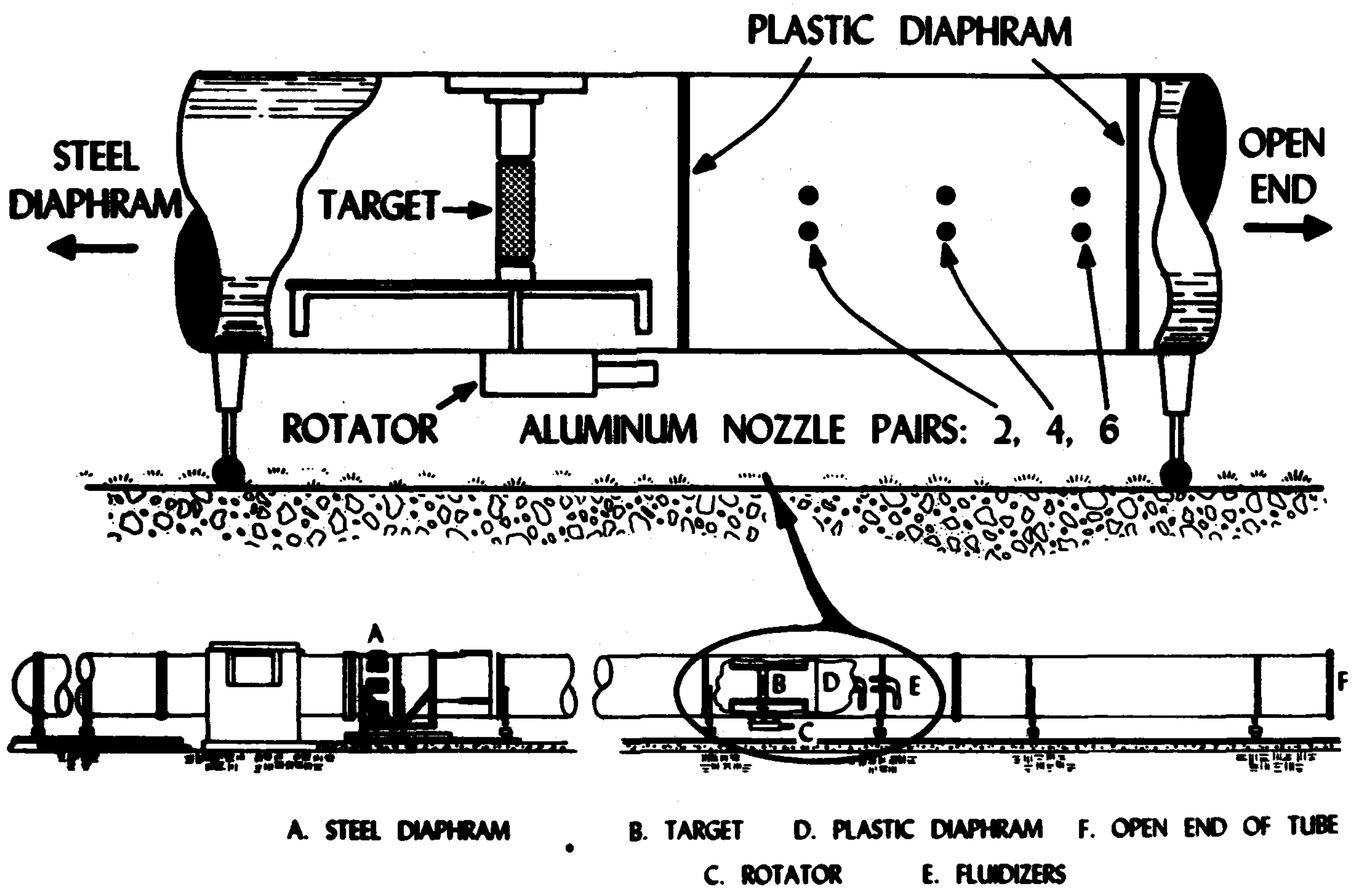

PLASTIC DIAPHRAM
STEEL DIAPHRAM
TARGET
OPEN END
ROTATOR
ALUMINUM NOZZLE PAIRS: 2, 4, 6
A
B
C
D
E
F
A. STEEL DIAPHRAM
B. TARGET
C. ROTATOR
D. PLASTIC DIAPHRAM
E. FLUIDIZERS
F. OPEN END OF TUBE

Figure 1. Ballistic Research Laboratory Thermal/Blast Simulator Facility

perforating the steel diaphragm at the interface of the driver section and the expansion section (Item A in Figure 1). Despite the fact that the thermal source was located downstream from the target, some degradation of the leading spike of the shock wave occurred due to the presence of heated air in the vicinity of the cylinder.

The following measurements were made on the cylinder during the tests: pressure, axial and circumferential strain, temperature, and thermal flux. In addition, two high speed cameras were used to monitor the sequence of events occurring in the simulator. Pre- and post-shot measurements of the radius of the cylinders were made at 384 points on each cylinder. More detailed descriptions of the instrumentation are contained in Reference 3.

The cylindrical targets were made from 1.016 mm thick sheets of aluminum 6061-T6. The sheets were formed into cylinders and seam welded on the side farthest from the loading side. The inside diameter of the cylinders was 30.48 cm (12 inches) and the length between upper and lower clamps was 80 cm. In order to make the ADINA finite element modeling of the experiments easier, the upper and lower ends of the cylinders were clamped securely in heavy steel end discs. The end discs were connected by shafts to bearings. The bearings were mounted in two heavy bearing mounts. To minimize the disturbance of the flow past the cylinder, the mounts were made with the same outer diameter as the target cylinders.

NUMERICAL MODEL

The numerical modeling of these tests was done with the ADINA finite element program [4]. The element type used was the 3-D brick element, which has a thermo-elastic formulation. Symmetry was assumed about a plane passing through the cylinder axis oriented in the direction of air-heat flow. Another plane of symmetry was assumed passing orthogonal to both the previous plane and passing through the Z-coordinate L/2, where L is the active length of the cylinder. With these assumptions, only one quarter of the cylinder was required for the finite element model. Actually, there was approximately a 10 per cent variation in heating along the axis of the cylinder due to nearby reflecting surfaces. However, this variation was ignored in the numerical simulation.

The quarter cylinder was modeled with 72 brick elements, 8 in the axial direction, 9 in the circumferential direction, and 1 in the thickness direction. The elements had a midside node on each edge and the number of integration points in the element was 3 in each direction, for a total of 27 per element. The outside surface of the undeformed mesh is indicated in Figure 2 by the crosses and dotted lines. The crosses represent the nodal points on the outside surface and the dotted lines are the undeformed mesh boundaries.

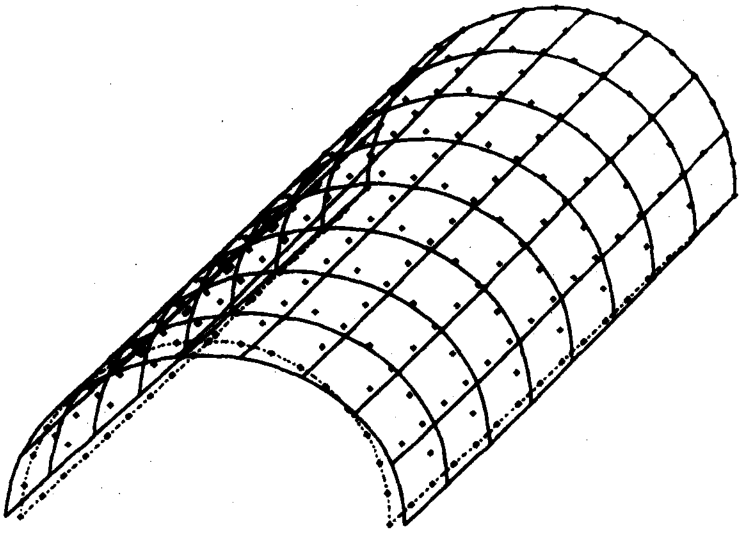


Figure 2. Mode Shape for the 7th Eigenfrequency (3961 cycles/seconds)

The lower eigenfrequencies and mode shapes were evaluated for use in a mode superposition analysis of the elastic response of the cold cylinder. Knowing the periods of the lower eigenfrequencies is also useful in determining the approximate time required for a peak deflection to occur. The periods determined ranged from 3.66 ms for the lowest mode to 1.47 ms for the eleventh mode. Shown in Figure 2 is mode number seven, which has the simpliest geometric shape (egg-shaped cross section).

The finite element calculations for the dynamic response of the cylinder were made using the explicit central difference time marching scheme. The time step was obtained from the Courant stability condition,

$$\Delta t = \frac{\Delta \ell}{\sqrt{E_y/\rho}}$$

where $\Delta \ell$ is the distance between the two closest nodes, E_y is Young's modulus, and ρ is the material density. For the higher order elements used here, we found that this time increment was unstable and we used one half of the Courant Δt, which was stable. This time increment was 50 ns.

The constitutive properties of the cold aluminum 6061-T6 material were measured in an Instron testing machine. Tensile test specimens were cut from the same sheet of material used for fabrication of the cylinders. Properties were measured in the rolling direction and perpendicular to the rolling direction. The average of these tensile properties was used in the ADINA calculations. Properties measured in this way were the Young's modulus, bilinear corner yield stress, plastic modulus and Poisson's ratio.

Dynamic temperature dependent constitutive properties were measured in the DNA Tri-Service Flash Test Facility. In these tests, samples of the material were heated in a fraction of a second and strained through the yield point several seconds later. These times were close to the temperature-time sequence experienced in the thermal/blast simulation tests. The yield points determined in these tests were consistent with other data [5]. However, the plastic tangent moduli appeared to be higher than expected. Due to this result and to difficulty in keeping strain gages bonded at the high temperatures, we chose to await further confirmation prior to using the data. The following temperature dependent stress-strain data have been derived from Reference 5. No temperature dependent data for the plastic tangent modulus and Poisson's ratio were available. The cold properties measured in the Instron testing machine of 650 MPa and 0.3285, respectively, were used for all temperature ranges.

Temperature °C	Density g/cm^3	Yield MPa	Young's Modulus GPa
28	2.703	301.0	64.73
100	2.689	275.4	62.20
150	2.678	259.1	60.80
200	2.667	240.5	59.29

The ADINA calculations were performed using a bilinear von Mises kinematic hardening material model. This model required the specification of the materials properties: Young's modulus, Poisson's ratio, initial yield stress, and plastic tangent modulus.

RESULTS

The three types of tests and analyses will be discussed in the same order as the experiments were performed. The input head-on pressure for the blast-only test is shown in Figure 3. The maximum stagnation overpressure was 60 kPa. Shown in Figure 4 is the short term reflected pressure on the cylinder at zero degrees. The peak load is 104 kPa (15 psi). Figure 5 shows the pressure loading on the side of the cylinder (90°). Loading data were recorded at 15° intervals around the cylinder. Since the finite element calculations required loading at 20° intervals around the circumference starting at 10°, these data were interpolated in both circumferential angle and time to supply the pressure loading for ADINA.

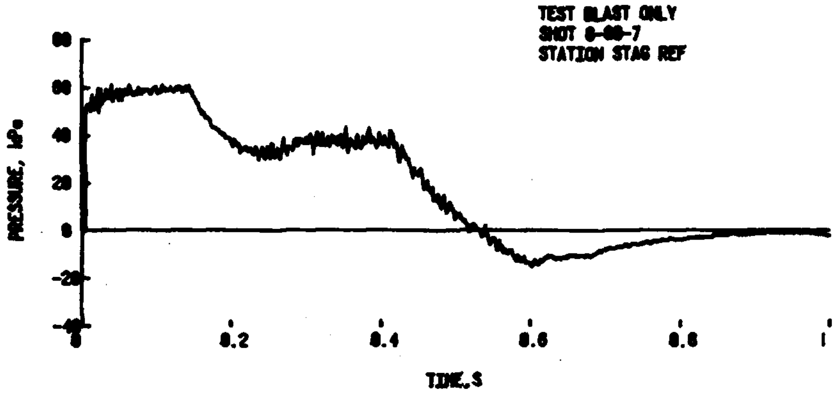


Figure 3. Blast-Only Test, Head-On Pressure

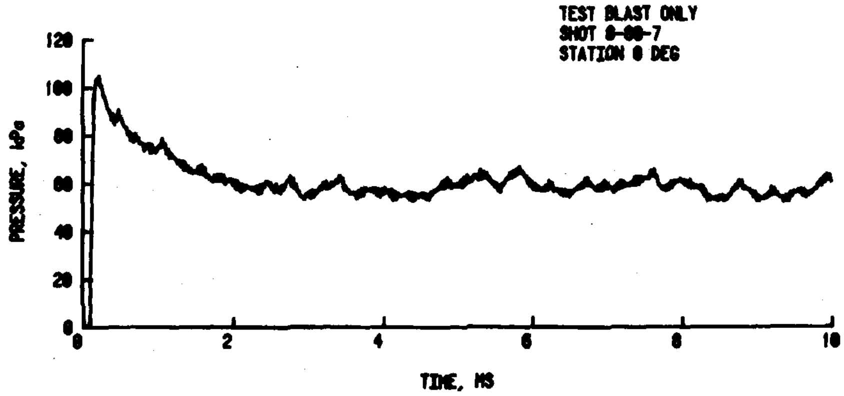


Figure 4. Blast-Only Test, Diffraction Phase Loading at 0°

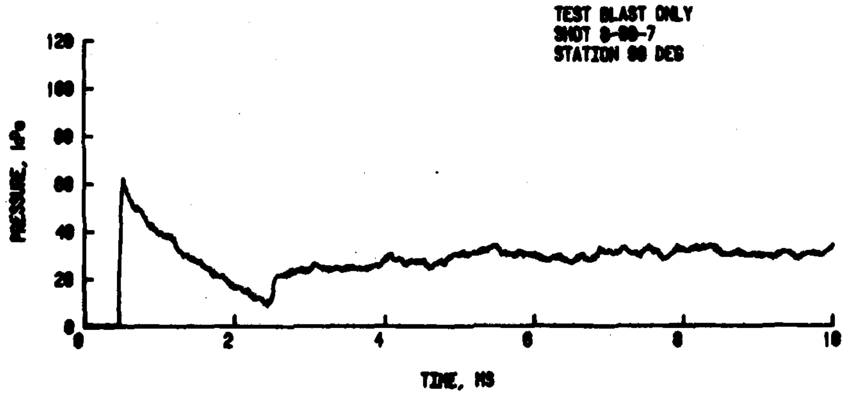


Figure 5. Blast-Only Test, Diffraction Phase Loading at 90°

Experimental measurements of the final radial deflection of the cylinder to the blast load were made and the results are shown in the top and middle part of Figure 6 for the right and left halves of the cylinder. The deflections shown have been magnified by a factor of ten to make the deformed shape more apparent. Deformation of the actual cylinder was barely noticeable to the eye. The resulting deformation pattern was only approximately symmetric about the zero degree line and about a plane at $Z = L/2$. For Figure 6, the loading direction is from the top for all three parts of the figure. The deformed shapes for the three subplots at the $Z = L/2$ center cross section are shown on the left side of Figure 6. The ADINA calculations have been carried out to 2.0 ms and the deflection pattern at that time is shown on the right side of Figure 6. Since no plasticity was observed in this calculation, this is not the final predicted shape of the cylinder, but is approximately the maximum deflection occurring during the excitation of the simulated cylinder.

In the second test, the cylinder was exposed to a thermal load only. Shown in Figure 7 is the flux and the fluence received by the flux gage positioned off to the side of the cylinder. The solid line is the flux as recorded by the gage. The fluence was obtained by integrating the flux curve. The nominal 25 cal/cm^2 fluence on the target was obtained by relating the fluence to previous calibration tests. Figures 8 and 9 show the temperature rise indicated by thermocouples at $Z = L/2$, $\theta = 0°$ and 90° respectively. The nominal temperature rise was 225°C at 0° and 45°C at 90°.

The ADINA model for the thermal-only loading was the same mesh configuration as for the blast-only loading. However, instead of the pressure load applied to the exterior of the cylinder, an initial displacement of the Z-coordinates of the cylinder proportional to the temperature increase was assumed. A prescribed displacement was applied to the symmetry end of the cylinder at $Z = L/2$. This artifice was used in lieu of performing an ADINAT [6] calculation to determine the temperature in the cylinder as a function of time. Earlier calculations using ADINAT showed that the heat penetrates through the thickness of 1.016 mm in a time which is short compared to the structural response time. Based on this result, temperature was assumed to reach equilibrium through the thickness instantaneously. Initial displacements were calculated for each nodal point based on the fluence received at that point. The displacements were calculated relative to the fixed end according to the formula

$$w(\theta,Z) = \frac{Z\alpha}{h\rho C} \int_0^\infty \Phi(\theta,t)\, dt,$$

where α is the coefficient of thermal expansion, h is the thickness, ρ is the density, C is the heat capacity, and $\Phi(\theta,t)$ is the incident flux.

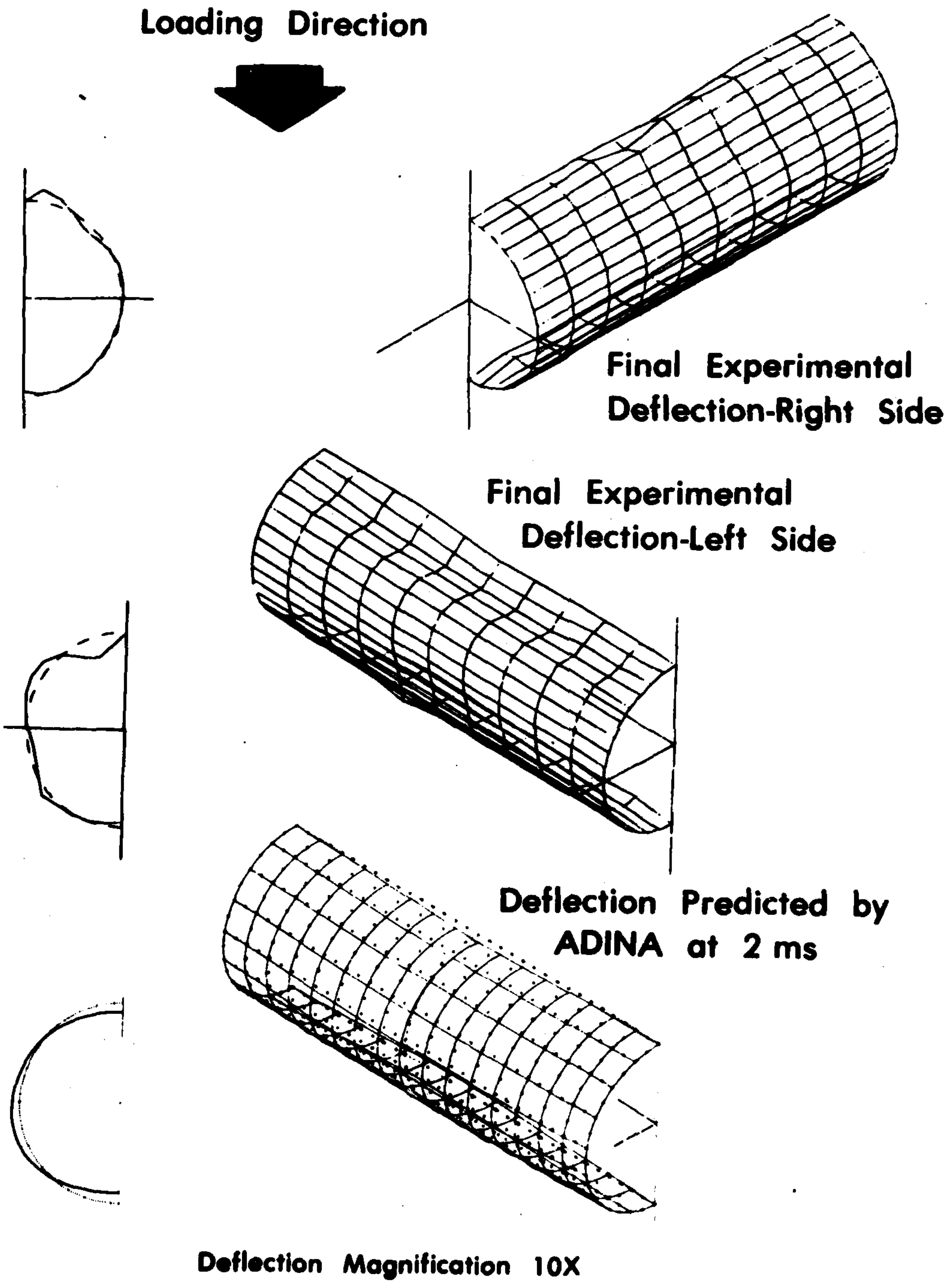

Loading Direction
Final Experimental
Deflection-Right Side
Final Experimental
Deflection-Left Side
Deflection Predicted by
ADINA at 2 ms
Deflection Magnification 10X

Figure 6. Comparison of Experimental and Predicted Deflections, Magnification 10X, Blast-Only Loading

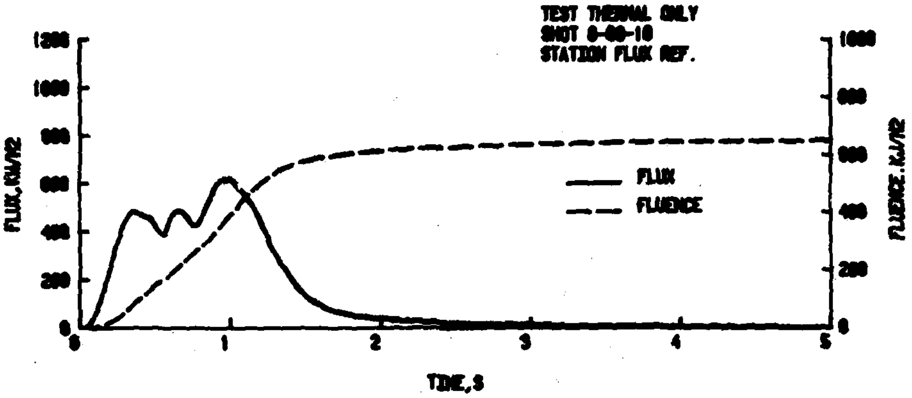


Figure 7. Reference Flux and Fluence

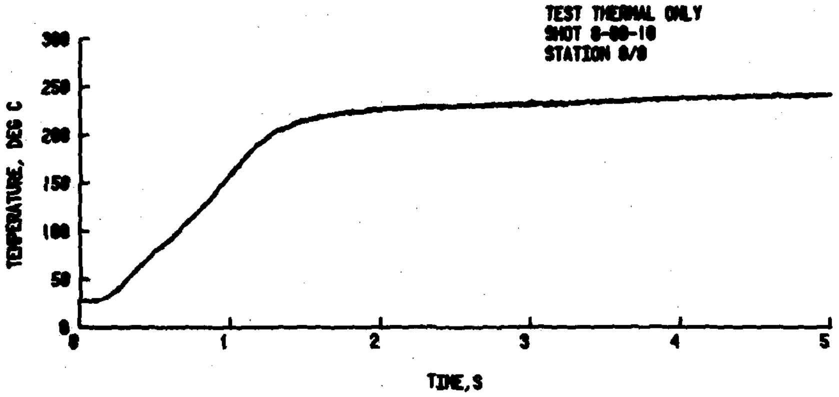


Figure 8. Temperature History at Z = L/2, θ = 0°

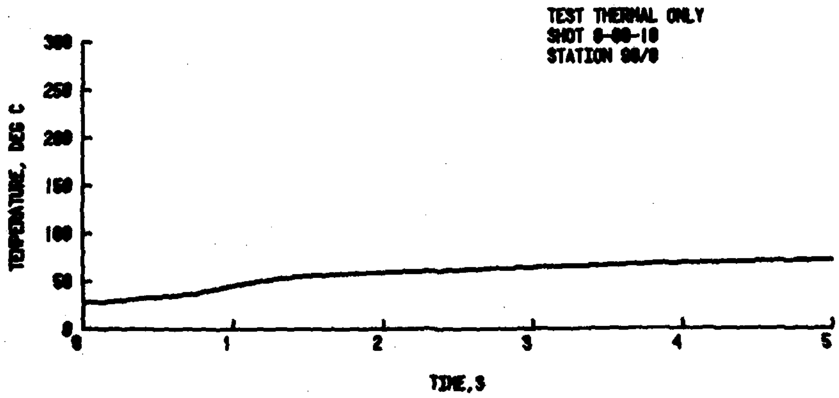


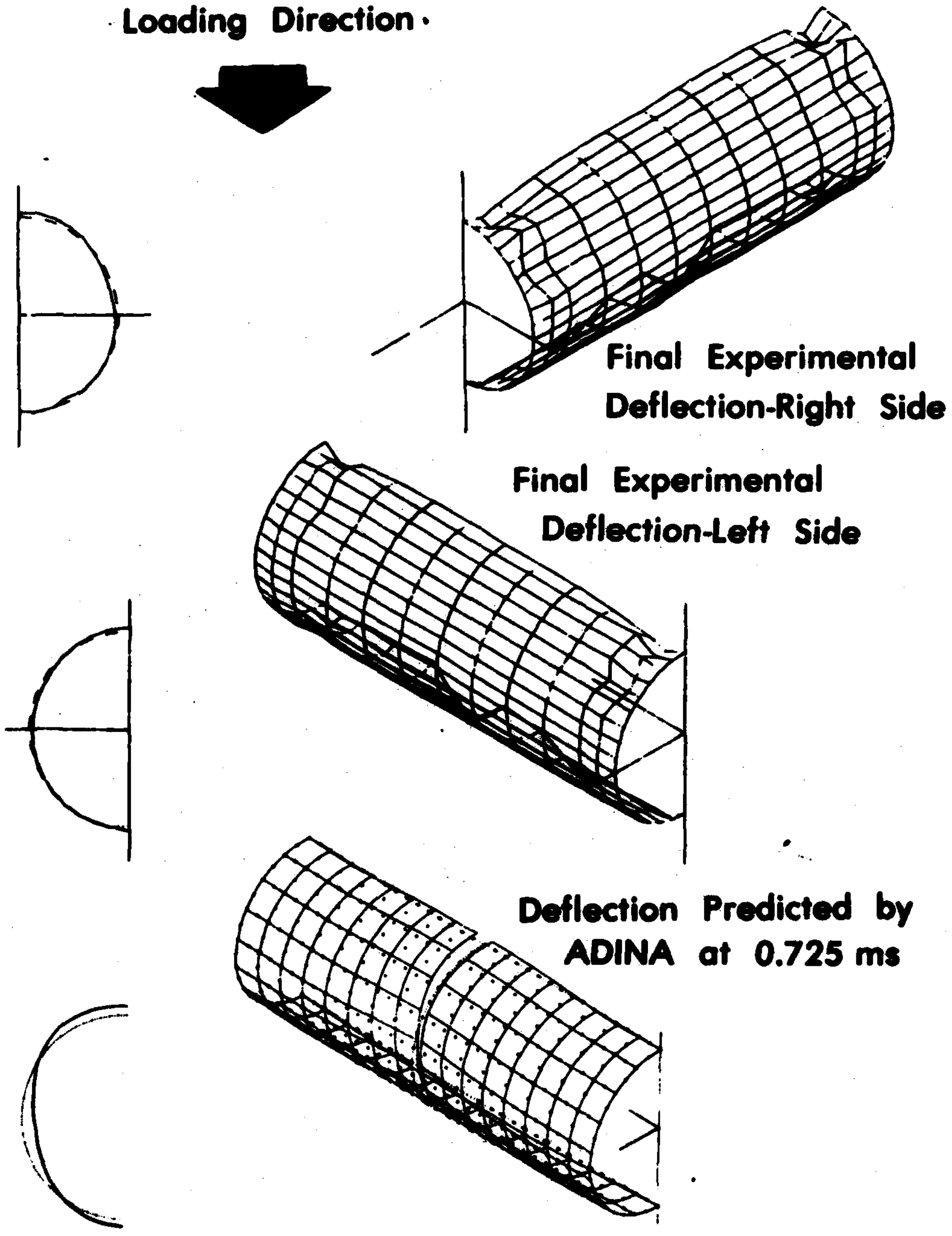
Figure 9. Temperature History at Z = L/2, θ = 90° for Thermal-Only Loaded Cylinder

Measurements on the post-test thermally loaded cylinder showed a slight degree of buckling near the ends of the cylinder on the near side. The deformation pattern is shown in the top two subplots of Figure 10 and is nearly symmetric in both the Z and θ coordinates. The displacements from the original geometric positions are magnified by 10X in Figure 10. The deflection was small and was most apparent to the unaided eye near each end where the largest changes in slope occurred. The deformation cross sections at the mid-point, $Z = L/2$, are shown on the left side of Figure 10. The displacement was outward at $\theta = 0°$ and inward at $\theta = 90°$.

The ADINA simulation of the thermally loaded cylinder was carried out to 1.1 ms. Plasticity occurred early in the calculation and peak deflections were observed at approximately 0.725 ms. The deformed shape at this time as predicted by ADINA is shown in the lower part of Figure 10. The apparent cut shown in Figure 10 at $Z = L/2$ is the prescribed constant Z-displacement which simulates the thermal expansion. The prescribed displacement ranged from 1.8 mm at $\theta = 0°$ to zero at $\theta = 120°$. The deformation cross section at $Z = L/2$ showed a similar trend as did the experimental results. However, no buckling was predicted by ADINA near the $Z = L/20$ point, contrary to what was observed in the experiment. An incipient buckling deflection of approximately twice the cylinder thickness appeared to be forming at $Z = L/2$, $\theta = 100°$ in the ADINA results.

In the final test in the BRL Thermal/Blast Simulator, the cylinder was exposed to the thermal radiation, rotated and exposed to the air shock. The reference flux and temperature histories at 0° and 90° are shown in Figures 11-13. The long term stagnation pressure record is shown in Figure 14. The pressure loadings on the cylinder at $\theta = 0°$ and 90° are shown in Figure 15 and 16. Comparing Figures 15 and 16 with Figures 4 and 5, one notices that the pressure load rises much more gradually in the presence of the heated air surrounding the cylinder. From Figures 14 and 3, one sees that the long term pressure produced by the shock tube is only slightly less for the thermal/blast test than was the case for the blast-only test. As shown in Figures 12 and 8, the peak temperature rise in thermal/blast test was 170°C and in the thermal-only test it was 225°C.

The final deformed shape of the cylinder after the combined thermal/blast test is shown in Figure 17. The deformation is symmetric with respect to the two symmetry planes. The deflection magnification in Figure 17 is only 5X, one half of the magnification used in Figures 6 and 10. At $Z = L/2$, the peak radial deflection for the thermal/blast test is approximately 4 times as large as for the blast-only deflection and 10 times as large as the thermal-only deflection. ADINA calculations for the thermal/blast simulation have not been made and no comparisons are available.


Loading Direction
Final Experimental
Deflection-Right Side
Final Experimental
Deflection-Left Side
Deflection Predicted by
ADINA at 0.725 ms

Deflection Magnification 10X

Figure 10. Comparison of Experimental and Predicted Deflections, Magnified 10X, Thermal-Only Loading

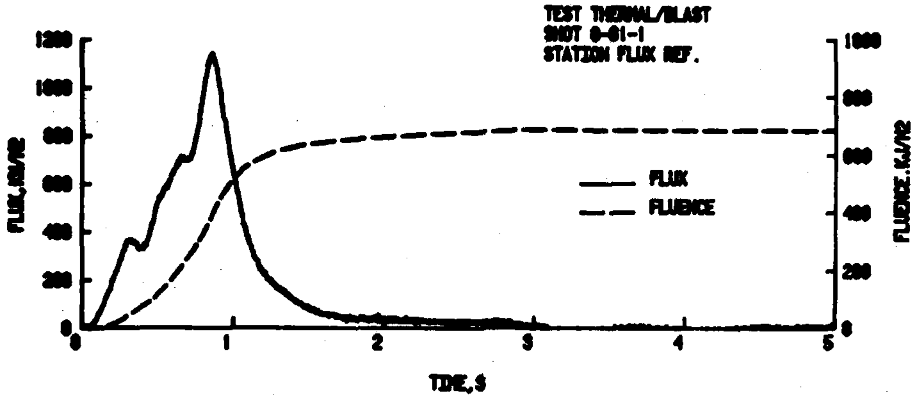


Figure 11. Thermal/Blast Test, Reference Flux and Fluence

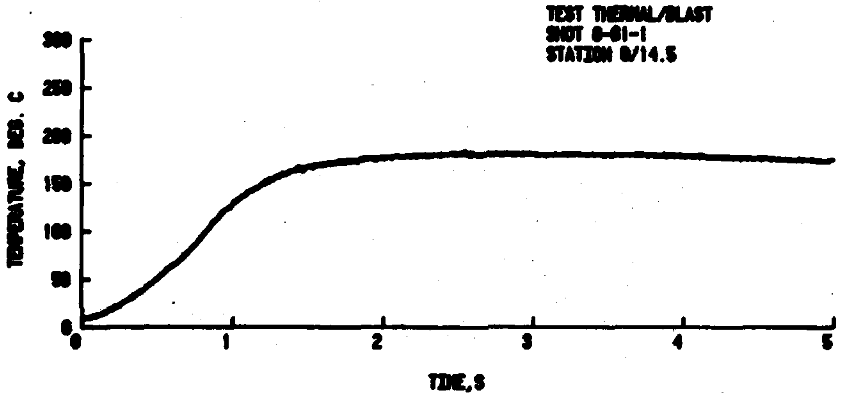


Figure 12. Thermal/Blast Test, Temperature History at Z = L/10, θ = 0°

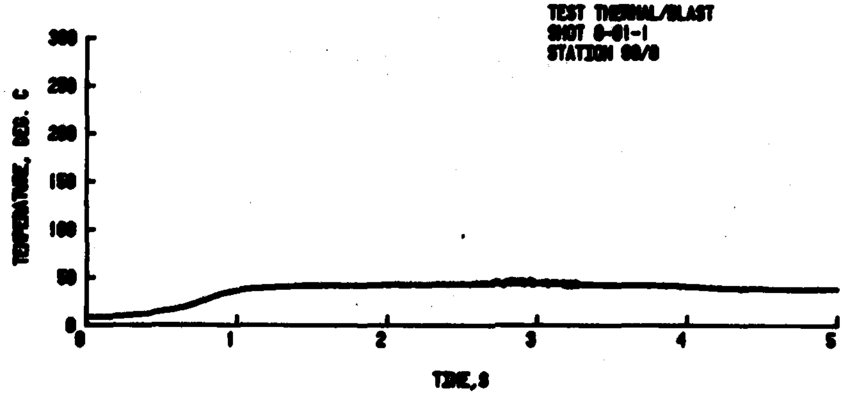


Figure 13. Thermal/Blast Test, Temperature History at Z = L/2, θ = 90°

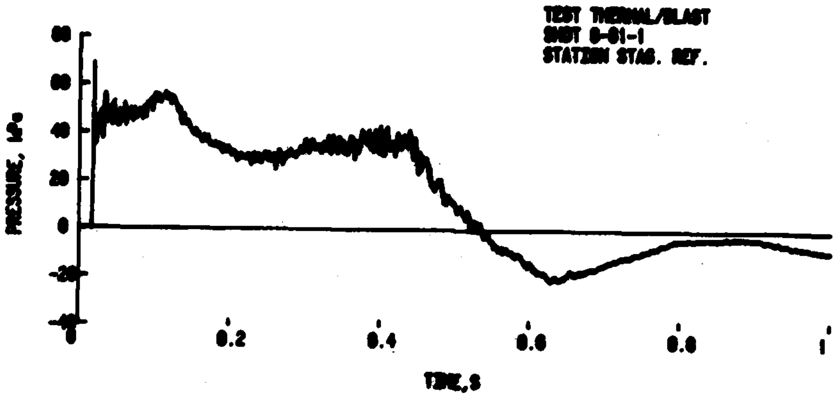


Figure 14. Thermal/Blast Test, Head-On Pressure

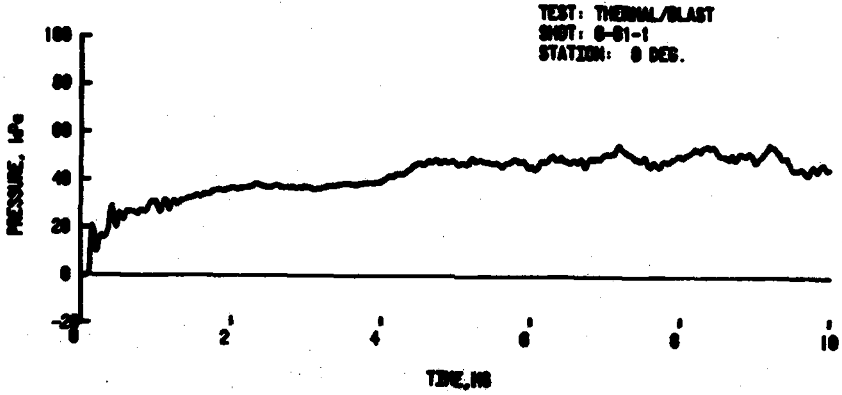


Figure 15. Thermal/Blast Test, Diffraction Phase Loading at 0°

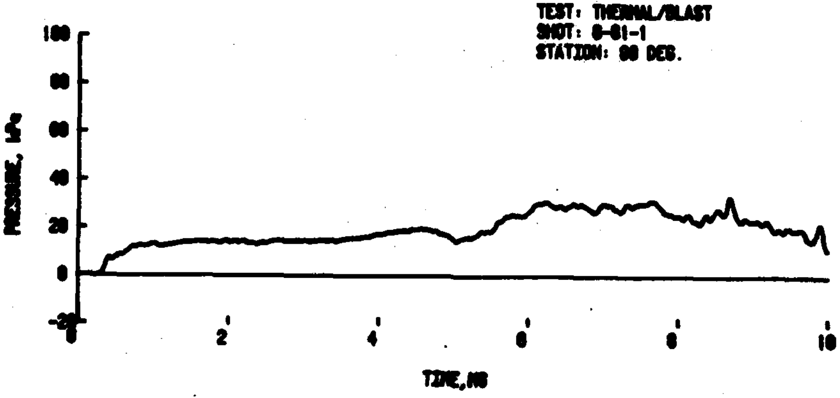
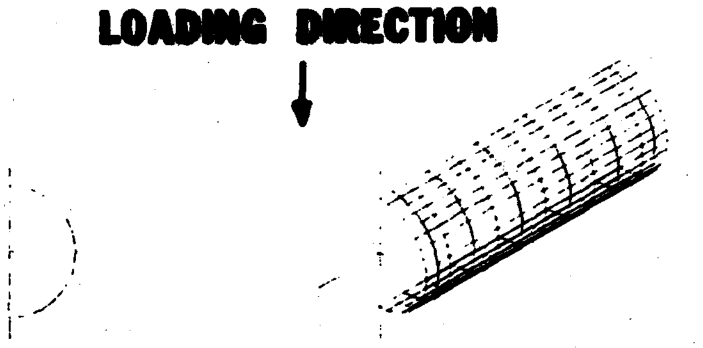
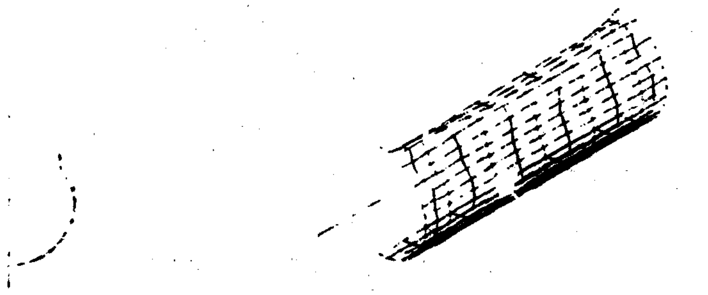


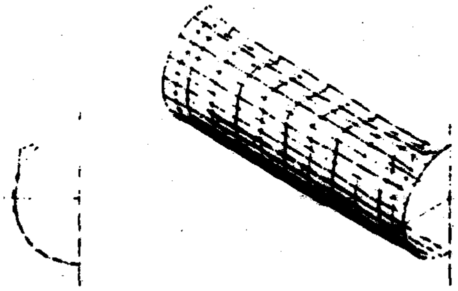
Figure 16. Thermal/Blast Test, Diffraction Phase Loading at 90°



PRE-TEST TARGET SHAPE - RIGHT HALF



POST-TEST TARGET SHAPE - RIGHT HALF



POST-TEST TARGET SHAPE - LEFT HALF

DEFLECTION MAGNIFICATION 5X

Figure 17. Comparison of Pre-Test and Post-Test Shapes for Thermal/Blast Test

CONCLUSIONS

The series of tests showed that the effect of combined thermal and blast loadings at realistic time separations produces a nonlinear response in thermally thin targets. Moreover, the shape of the deformed cylinders is markedly different between the three tests. The heated cylinder tended to displace outward on the near and far sides; whereas, the combined test resulted in a dished-in surface on the near side.

The results obtained thus far with ADINA are not as satisfying as was originally expected. However, the permanent deflections produced in both the thermal-only and blast-only tests were very small. Furthermore, the general shape of the deformations predicted by ADINA was correct. It is evident that we need to develop a Kirchhoff shell element whose stable time step is not limited by the rotatory inertia. This would improve the efficiency of the ADINA calculation significantly.

REFERENCES

1. P. J. Morris, "A Review of Research on Thermal Radiation Degradation of Structural Resistance to Air Blast," (U), URS Research Company, DNA-2856F, December 1971 (SECRET).

2. D. M. Wilson, "The Distribution and History of Temperature in Circular Cylinders Exposed to the Thermal Radiation Pulse of a Nuclear Detonation," NOLTR-71-61, June 1971.

3. R. J. Pearson, et al, "Synergism in Nuclear Thermal/Blast Loading," Seventh International Symposium on Military Applications of Blast Simulation, Medicine Hat, Alberta, Canada, July 1981.

4. K. J. Bathe, "ADINA, A Finite Element Program for Automatic Dynamic Incremental Analysis," MIT-82448-1, December 1978.

5. T. Stefansky, et al, "Temperature Induced Degradation of Mechanical Properties Following Instantaneous Heating," AFWL-TR-71-62, Sep 1977.

6. K. J. Bathe, "ADINAT, A Finite Element Program for the Automatic Dynamic Incremental Nonlinear Analysis of Temperatures," MIT-82448-5, Dec. 1978.

*U.S. GOVERNMENT PRINTING OFFICE: 1982--342-007/839